

Search for characteristic structural features of mammalian mitochondrial tRNAs

MARK HELM,^{1,3} HERVÉ BRULÉ,^{1,4} DAGMAR FRIEDE,² RICHARD GIEGÉ,¹
DOERN PÜTZ,¹ and CATHERINE FLORENTZ¹

¹Unité Propre de Recherche 9002 du Centre National de la Recherche Scientifique,
Département Mécanismes et Macromolécules de la Synthèse Protéique, et Cristallogénèse,
Institut de Biologie Moléculaire et Cellulaire, F-67084 Strasbourg Cedex, France

²Institute of Theoretical Chemistry, A-1090 Vienna, Austria

ABSTRACT

A number of mitochondrial (mt) tRNAs have strong structural deviations from the classical tRNA cloverleaf secondary structure and from the conventional L-shaped tertiary structure. As a consequence, there is a general trend to consider *all* mitochondrial tRNAs as “bizarre” tRNAs. Here, a large sequence comparison of the 22 tRNA genes within 31 fully sequenced mammalian mt genomes has been performed to define the structural characteristics of this specific group of tRNAs. Vertical alignments define the degree of conservation/variability of primary sequences and secondary structures and search for potential tertiary interactions within each of the 22 families. Further horizontal alignments ascertain that, with the exception of serine-specific tRNAs, mammalian mt tRNAs do fold into cloverleaf structures with mostly classical features. However, deviations exist and concern large variations in size of the D- and T-loops. The predominant absence of the conserved nucleotides G18G19 and T54T55C56, respectively in these loops, suggests that classical tertiary interactions between both domains do not take place. Classification of the tRNA sequences according to their genomic origin (G-rich or G-poor DNA strand) highlight specific features such as richness/poorness in mismatches or G-T pairs in stems and extremely low G-content or C-content in the D- and T-loops. The resulting 22 “typical” mammalian mitochondrial sequences built up a phylogenetic basis for experimental structural and functional investigations. Moreover, they are expected to help in the evaluation of the possible impacts of those point mutations detected in human mitochondrial tRNA genes and correlated with pathologies.

Keywords: aminoacylation identity; bizarre tRNAs; compilation; G-U pairs; neurodegenerative disorders; T-loop

INTRODUCTION

Transfer RNAs (tRNAs) are key molecules in protein biosynthesis in all living organisms. Whereas in prokaryotes a single set of tRNAs is present in the cytosol, additional sets are present in and specific to the organelles in eukaryotes. To fulfill their biological role, tRNAs do have very specific structural properties that allow an optimal positioning of signals for interaction with various partners such as the cognate aminoacyl-tRNA synthetases (the enzymes that charge the correct amino acid

to the 3' end of the specific tRNAs), translational initiation or elongation factors, and the ribosomal machinery. Achievement of the structural properties themselves is linked to precise conserved or semiconserved sequence information, and recognition by maturation enzymes (e.g., RNase P, nucleotidyl transferase, post-transcriptional modification enzymes). Presently more than 5,000 individual tRNA sequences are known (Sprinzl et al., 1998), from which a majority can be folded into the “canonical” cloverleaf secondary structure, and further into a three-dimensional L-shaped form based on long-range tertiary interactions (Fig. 1A) (Söll & RajBhandary, 1995). For these tRNAs, mainly prokaryotic and cytosolic from eukaryotes, the important signals for interaction with partners have mostly been decrypted (see, e.g., Giegé et al., 1998; Grosjean & Benne, 1998).

Organellar tRNAs, however, are by far less well known and may deviate already at the structural level from the

Reprint requests to: Catherine Florentz, Unité Propre de Recherche 9002 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15, rue René Descartes, 67084 Strasbourg Cedex, France; e-mail: C.Florentz@ibmc.u-strasbg.fr.

³Present address: California Institute of Technology, Division of Biology, Pasadena, California 91125, USA.

⁴Present address: Institute for Structural Biology and Drug Discovery, Virginia Biotechnology Research Park, 800 East Leigh Street, Richmond, Virginia 23219, USA.

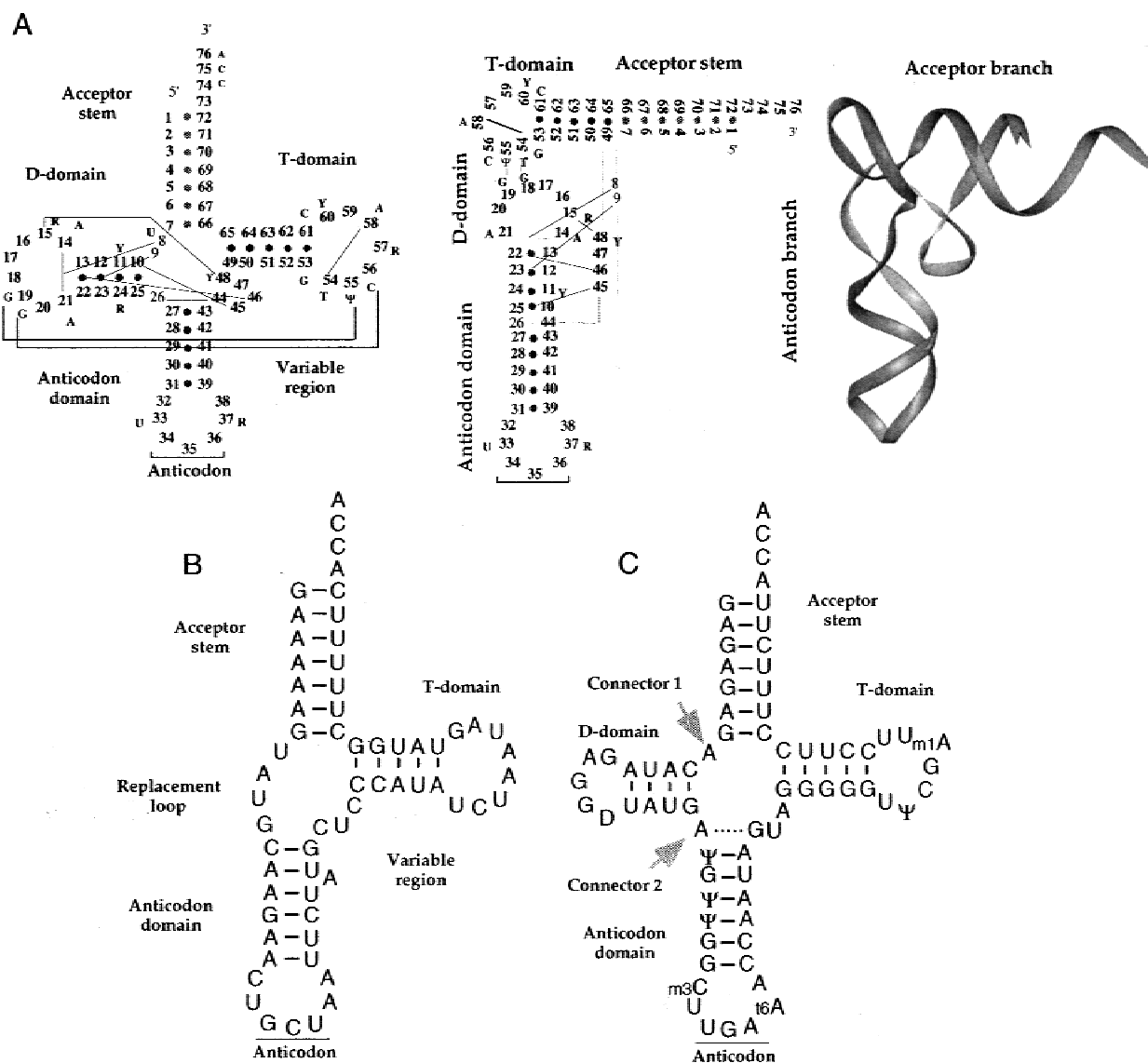


FIGURE 1. A: From secondary to tertiary structures of "classical" tRNAs. Positions of nucleotides are numbered according to conventional rules (Sprinzl et al., 1998). Conserved elements in the cloverleaf are highlighted as well as their involvement in folding of the L-shaped tertiary structure. Nomenclature is as follows: Y: pyrimidine; R: purine; Ψ : pseudouridine. B: The particular case of bovine mt tRNA^{Ser(UCN)} (Watanabe et al., 1994a; Hayashi et al., 1998). Notice the absence of the complete D-domain and its replacement by a connector. Also, the length of the anticodon stem and the size of the T-loop are at variance with classical tRNAs. C: The particular case of bovine mt tRNA^{Ser(AGY)} (Ueda et al., 1983; Hayashi et al., 1997b). The main differences from the classical cloverleaf reside in the length of connector 1, an additional pair in the anticodon stem, and the short size of the variable region.

canonical tRNAs (reviewed in Dirheimer et al., 1995). Moreover, knowledge about functional properties of these tRNAs remains limited. Since the discovery of mitochondrial (mt) tRNAs from nematodes that are missing complete domains of the cloverleaf structure (either the D-domain or the T-domain; see Fig. 1 for nomenclature of the different tRNA domains) and have instead short stretches of nucleotides referred to as replacement loops, mitochondrial tRNAs are referred to as "bizarre" tRNAs (Wolstenholme et al., 1987, 1994; Watanabe et al., 1994b). The absence of the complete D-arm was also reported for several mammalian mt

tRNA^{Ser(AGY)} (Fig. 1B; de Bruijn et al., 1980; Sprinzl et al., 1998). Apparently, more moderate deviations from the canonical cloverleaf, including changes in the number of nucleotides at the connectors as well as shortening of the D- and/or T-stems and loops, and/or elongated anticodon stems, have also been reported (Steinberg & Cedergren, 1994; Steinberg et al., 1994; Fig. 1C). Structure and function of some such mt tRNAs have been investigated by sequence comparison, structural probing, UV melting, and NMR experiments, and computer modeling has permitted investigators to propose three-dimensional structures (see, e.g., de Bruijn

et al., 1980; de Bruijn & Klug, 1983; Watanabe et al., 1994a; Hayashi et al., 1997b, 1998). Efforts to explain other similar deviations from the cloverleaf structure have focused on a computer-model-based structural rational (Steinberg et al., 1997). Whereas the above discussed mt tRNAs, picked from various organisms, indeed show significant deviations from the classical cloverleaf structure, it remains to be analyzed whether *all* mt tRNAs deviate as well, and if so, to evaluate the degree of deviation.

Mammalian mt tRNAs form a family of molecules for which poor structural and functional knowledge is available. This is a considerable drawback, especially at a time when a growing number of human diseases are found to be correlated to point mutations in mt tRNA genes (Wallace, 1992, 1999; Larsson & Clayton, 1995; Schon et al., 1997). Since the first description in 1990 (Goto et al., 1990) of a correlation between the MELAS syndrome and point mutation at position 3243 in the tRNA^{Leu(UUR)} gene, more than 70 mutations have been reported so far in 19 out of the 22 tRNA genes (Kogelnick et al., 1998, and updated web site Mitomap).

Understanding the functional properties of mammalian mt tRNAs requires an understanding of their underlying structural basis. Thus, we present here a sequence alignment of the 22 tRNA genes in 31 fully sequenced mammalian mt genomes. This survey has as a primary goal to decipher the differences and similarities to canonical tRNAs. Any conserved structural features in the mt tRNAs that deviate from the canonical tRNAs would be of particular interest, because they constitute landmarks to distinguish these two types of tRNAs from one another in structure. We have therefore paid special attention to the degree of evolutionary conservation of such deviations. In a number of previously reported sequence alignments, comparisons of mt tRNAs had been made in a "horizontal" way among the different tRNAs inside one species, that is, among tRNAs of different amino acid specificity and isoacceptors (Kumazawa & Nishida, 1993; De Giorgi et al., 1996; Lynch, 1996). These sequence alignments in general showed a very low degree of conservation. Alternatively, "vertical" alignments based on comparison of a number of tRNAs of a same specificity but among different species allow us to assess specific features for a given amino acid identity. Such alignments have so far suffered from the limited number of mt tRNA sequences of species in a class of organisms (Cantatore et al., 1982; Watanabe et al., 1994a; De Giorgi et al., 1996). With the plethora of mitochondrial genome sequences that have recently been published (Boore, 1999), there are now a great number of mammalian mt tRNA sequences available for systematic vertical alignment, representing a wealth of structural information that has so far gone unexploited.

By sequence alignments of 679 mammalian genomic tRNA sequences, we here identify conserved structural

features as well as domains that are subject to frequent mutation for each of the 22 tRNA specificities (it is recalled that in mitochondria there is one tRNA species for each amino acid, except for leucine and serine, for which there are two isoacceptor tRNAs). The examined secondary structure elements comprise lengths of loops, stems, and connectors, conservation of Watson-Crick and G-T base pairings, as well as a mismatch compilation. Tertiary interactions were also sought on the basis of structural knowledge of classical tRNAs and on that of bovine mt tRNA^{Ser(AGY)} (Ueda et al., 1983; Hayashi et al., 1997b), tRNA^{Ser(UCN)} (Watanabe et al., 1994a; Hayashi et al., 1998), and tRNA^{Phe} (Wakita et al., 1994). Finally, conservation in primary sequences for each tRNA specificity were related to functional knowledge compiled on canonical tRNAs (Giegé et al., 1993, 1998; Söll & RajBhandary, 1995).

It is expected that the present results will be a useful basis for structural and functional investigations of mammalian mt tRNAs. The determination of crucial elements for interaction with molecular partners will be facilitated at a time when more and more tRNA-interacting proteins from animal mitochondria are being identified (Tiranti et al., 1997) and cloned (Chihade et al., 1998; Takeuchi et al., 1998; Bullard et al., 1999, 2000). Moreover, this sequence compilation is expected to serve as a basis for the understanding of the impact of pathologic mutations in human mt tRNA genes (Schon et al., 1997) on the dysfunctioning of the corresponding tRNAs.

RESULTS

All mammalian mitochondrial genomes possess 22 tRNA genes, one specific for each amino acid, but two for leucine and serine, according to their codon-decoding abilities. These genes are transcribed either from the heavy DNA strand or from the light DNA strand [the heavy strand is G rich, the light strand is C rich (Anderson et al., 1981)]. The distribution of tRNA genes on the heavy and the light DNA strands, as well as their polarity (i.e., the order of appearance in the genome), are similar in all mammalian mt genomes (Boore, 1999, and references therein). Thus, transcription of the heavy DNA chains leads to 14 "light" tRNAs and transcription of the light DNA chains leads to 8 "heavy" tRNAs (with a single exception in *Didelphis virginiana*). We present here the compilation of the 22 tRNA species in 31 mammalian genomes (Table 1). Results obtained from manual alignment of sequences are presented in Figure 2 in the form of "typical tRNAs" and of "consensus tRNAs." Whereas typical tRNAs highlight the most likely primary, secondary, and tertiary features in each family, consensus tRNAs depict conserved and semiconserved nucleotides. The detailed distribution of nonclassical G-T base pairs and mismatches is also indicated in Figure 2, as well as the sizes of D- and T-loops. tRNA

TABLE 1. Mitochondrial genomes used in this study. For references see Boore (1999).

Taxonomy	Species (Latin)	Species (English)	Accession Number
Eutheria			
Primates	<i>Homo sapiens</i>	Human	V00662
	<i>Gorilla gorilla</i>	Gorilla	D38114
	<i>Pan paniscus</i>	Common Chimpanzee	D38116
	<i>Pan troglodytes</i>	Bonobo Chimpanzee	D38113
	<i>Pongo pygmaeus</i>	Orang Utan	D38115
	<i>Hylobates lar</i>	Gibbon	X99256
	<i>Papio hamadryas</i>	Baboon	Y18001
Perissodactylae	<i>Equus caballus</i>	Horse	X79547
	<i>Equus asinus</i>	Donkey	X97337
	<i>Rhinoceros unicornis</i>	Indian Rhinoceros	X97336
	<i>Ceratotherium simum</i>	White Rhinoceros	Y07726
Cetartiodactylae	<i>Bos taurus</i>	Cow	V00654
	<i>Ovis aries</i>	Sheep	AF010406
	<i>Sus scrofa</i>	Pig	AF034253
	<i>Balaenoptera musculus</i>	Blue Whale	X72204
	<i>Balaenoptera physalus</i>	Fin Whale	X61145
Hippopotamidae	<i>Hippopotamus amphibius</i>	Hippopotamus	AJ010957
Rodentia	<i>Mus musculus</i>	Mouse	V00711
	<i>Rattus norvegicus</i>	Rat	X14848
	<i>Myoxus glis</i>	Fat Dormouse	AJ001562
Lagomorpha	<i>Oryctolagus cuniculus</i>	Rabbit	AJ001588
Carnivora	<i>Felis catus</i>	Domestic Cat	U20753
	<i>Halichoerus grypus</i>	Gray Seal	X72004
	<i>Phoca vitulina</i>	Harbour Seal	X63726
	<i>Canis familiaris</i>	Dog	U96639
Chiroptera	<i>Artibeus jamaicensis</i>	Jamaican fruit-eating bat	AF061340
Insectivora	<i>Erinaceus europeus</i>	Hedgehog	X88898
Edentata	<i>Dasypus novemcinctus</i>	Nine-banded Armadillo	Y11832
Metatheria	<i>Didelphis virginiana</i>	North American Opossum	Z29573
	<i>Macropus robustus</i>	Wallaroo	Y10524
Monotremata	<i>Ornithorhynchus anatinus</i>	Duckbill Platypus	X83427

families are classified according to their genomic origin (heavy or light strand) and polarity, and are expressed as DNA sequences.

In what follows, the outstanding characteristic features of mammalian mt tRNAs are analyzed. Because mt tRNAs have often been referred to as bizarre molecules because of large differences in secondary structures when compared to classical tRNAs, an important prerequisite to individual nucleotide comparison in the present work was the search for secondary structure alignments. Thus, secondary features will be discussed first.

Secondary structure features in mammalian mitochondrial tRNAs

Global view

As seen in Figure 2, within the 22 analyzed tRNA families, 20 are constituted by tRNAs that do fold into classical cloverleaves, the 2 exceptions being the 2 tRNA^{Ser} families. tRNA^{Ser(AGY)} have the most deviating structures, as all are missing the D-arm. In the case of

tRNA^{Ser(UCN)}, the four branches of the cloverleaf are present, but there is only 1 nt at the level of the first connector, and the anticodon stem is formed by 6 bp instead of 5. For both tRNA^{Ser} families, these structural features, which have been identified and studied previously (de Bruijn & Klug, 1983; Hayashi et al., 1997a, 1997b, 1998), are conserved in the 31 animal species analyzed. All the 20 other tRNA families are composed of molecules that can fold into cloverleaves with main typical features (with the exception of one tRNA^{Cys} and two tRNA^{Lys}). They share a 7-bp acceptor stem, connector 1 formed of nt 8 and 9 (tRNA numbering), a D-stem and loop, a second connector formed by a single nucleotide, a 5-bp anticodon stem and a 7-membered anticodon loop, a variable region, and a T-stem and arm. The sizes of these domains are conserved for all families, with the noticeable exceptions of the D- and T-loops (see below). While most strongly conserved (either 4 or 5 nt) in 19 families, the variable region presents some small deviations within both tRNA^{Leu} families. The conserved small size of the variable regions in all the 20 tRNA families is in contrast to the situation in classical tRNAs where the class II tRNAs

A “Light” tRNAs

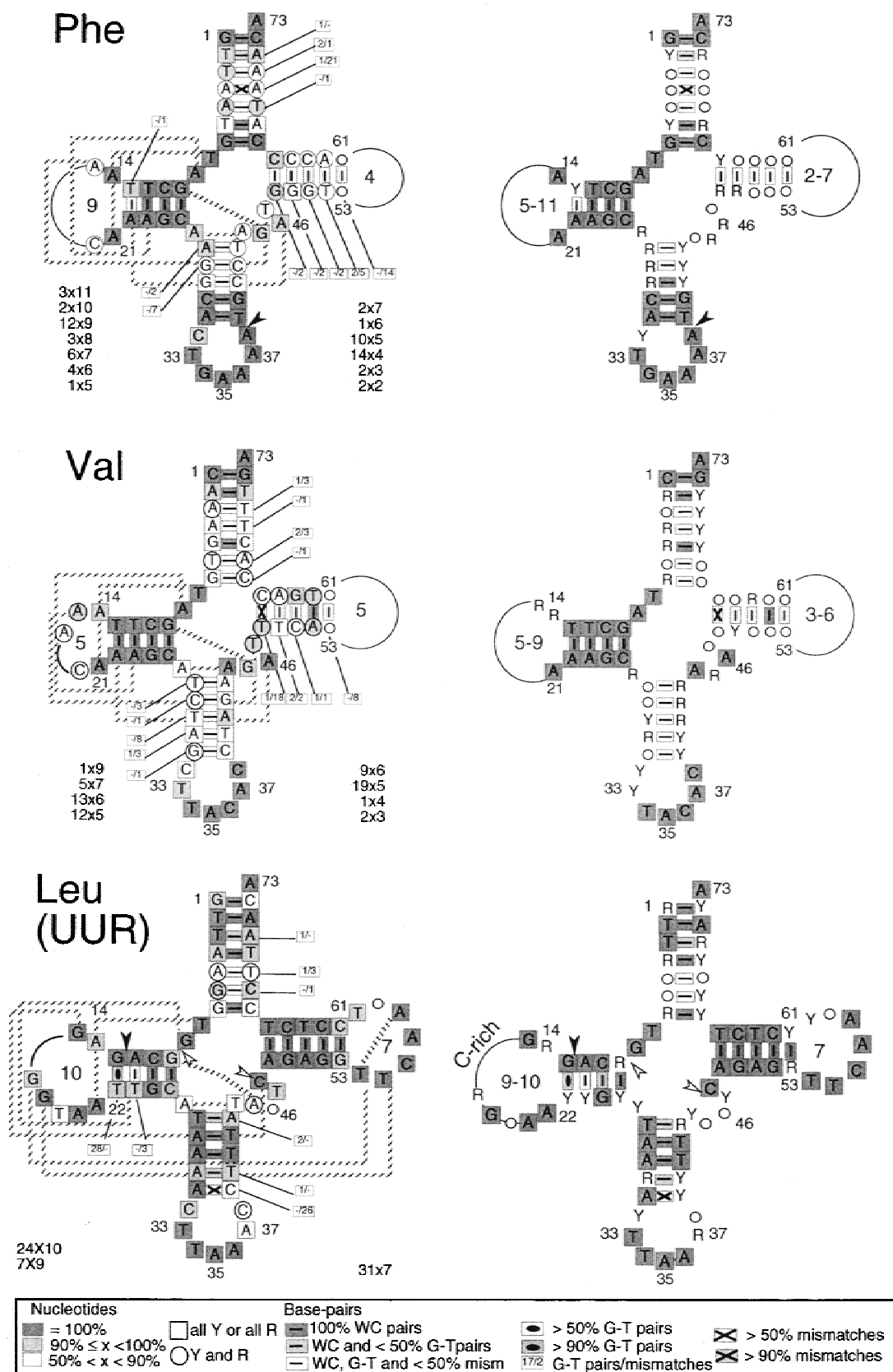


FIGURE 2. (Figure continues on next seven pages; legend is on page 1364.)

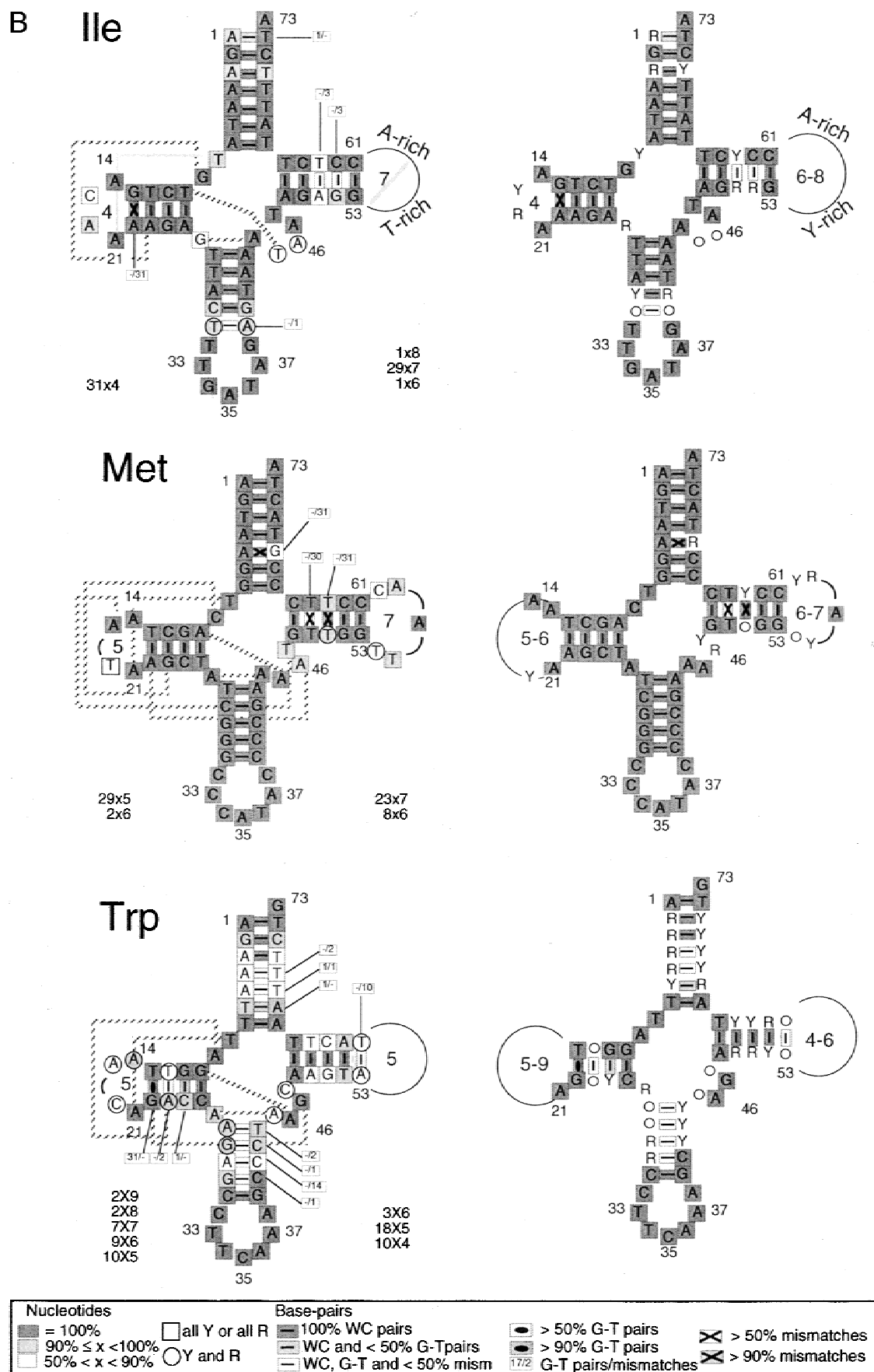
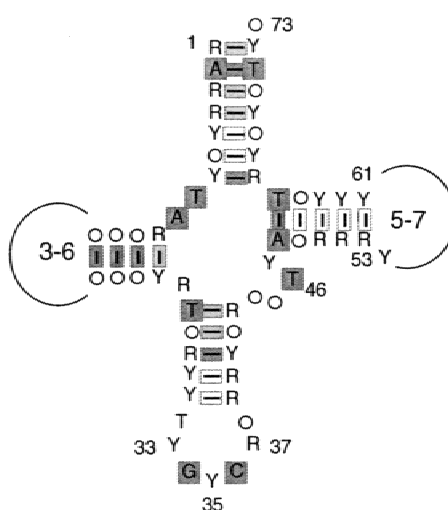
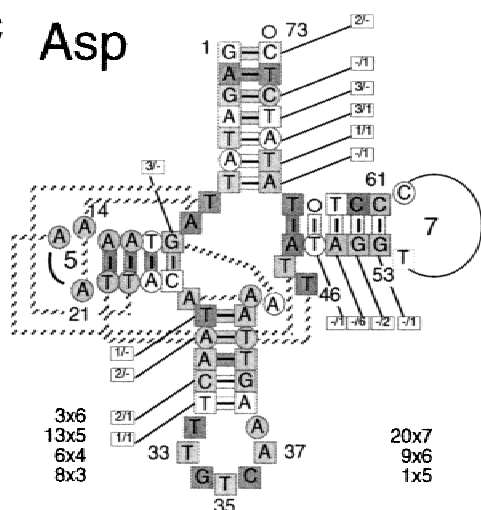
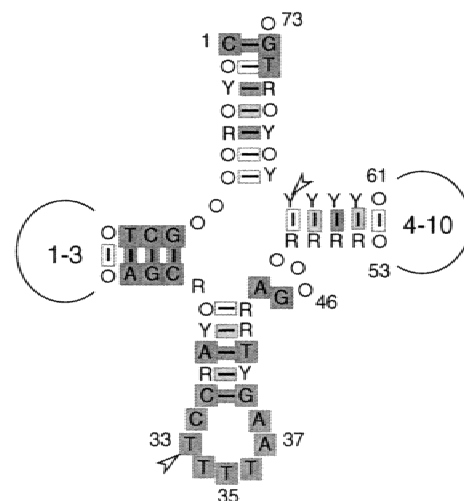
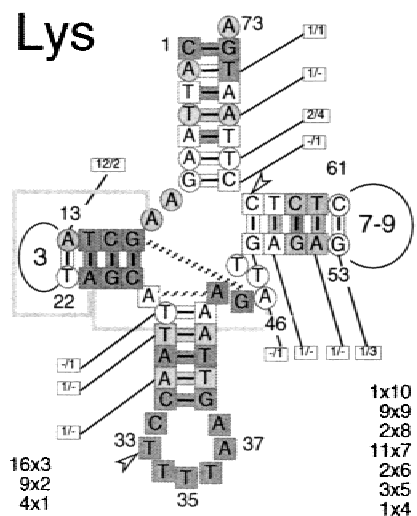


FIGURE 2. (Continued.)

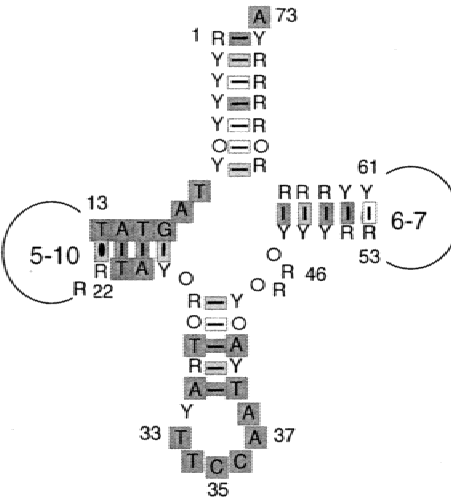
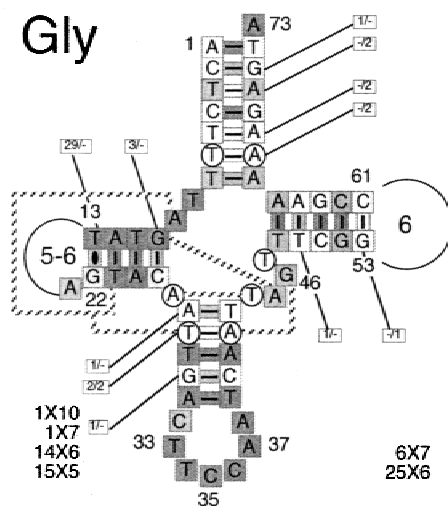
C Asp



Lys



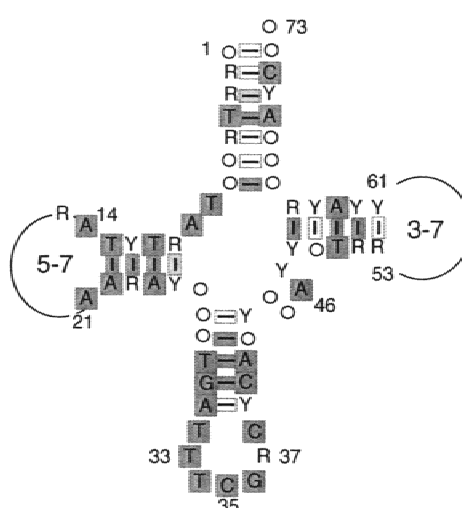
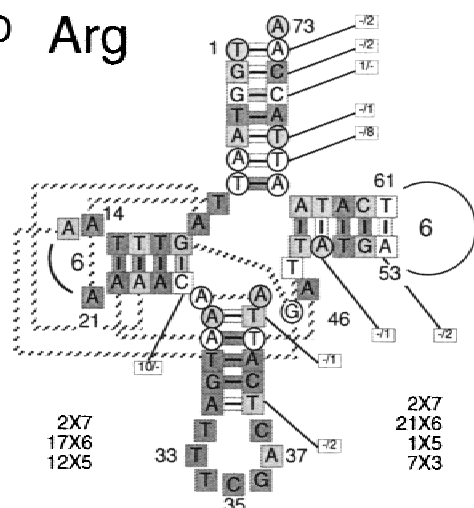
Gly



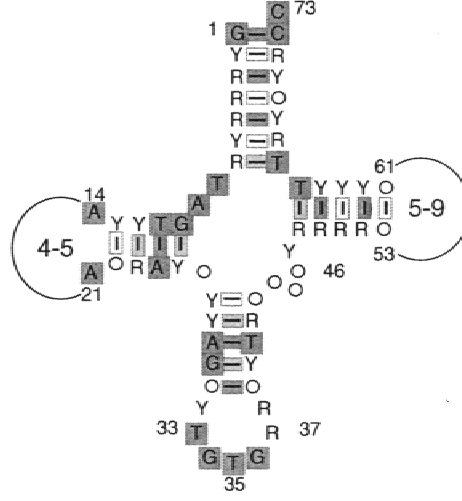
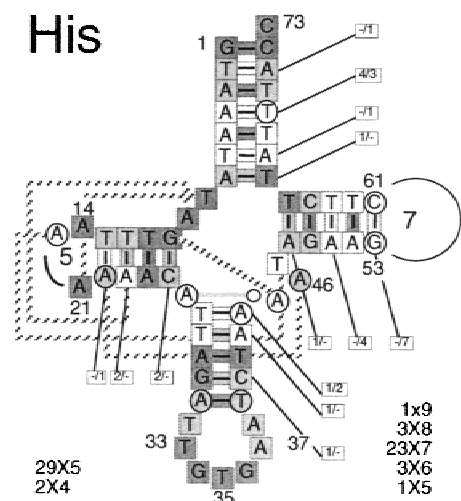
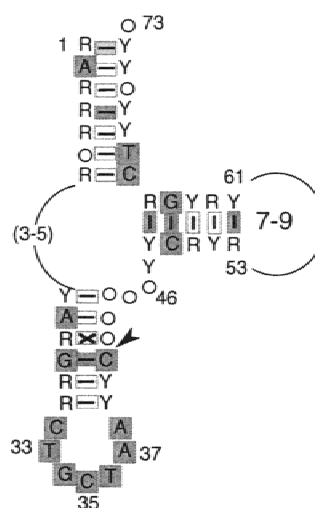
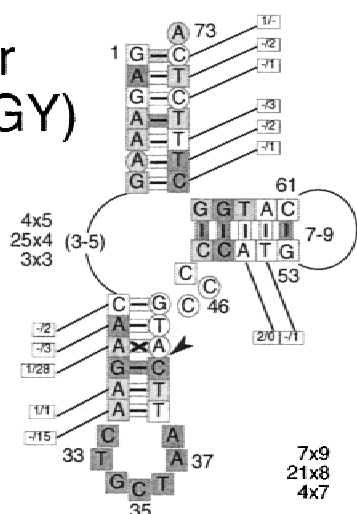
Nucleotides		Base-pairs	
	= 100%		all Y or all R
	90% ≤ x < 100%		100% WC pairs
	50% < x < 90%		WC and < 50% G-T pairs
	Y and R		WC, G-T and < 50% mism
			> 50% G-T pairs
			> 90% G-T pairs
			G-T pairs/mismatches
			> 50% mismatches
			> 90% mismatches

FIGURE 2. (Continued.)

D Arg



His

Ser
(AGY)

Nucleotides		Base-pairs	
= 100%	all Y or all R	100% WC pairs	> 50% G-T pairs
90% ≤ x < 100%	Y and R	WC and < 50% G-Tpairs	> 90% G-T pairs
50% < x < 90%		WC, G-T and < 50% mism	G-T pairs/mismatches
			> 50% mismatches
			> 90% mismatches

FIGURE 2. (Continued.)

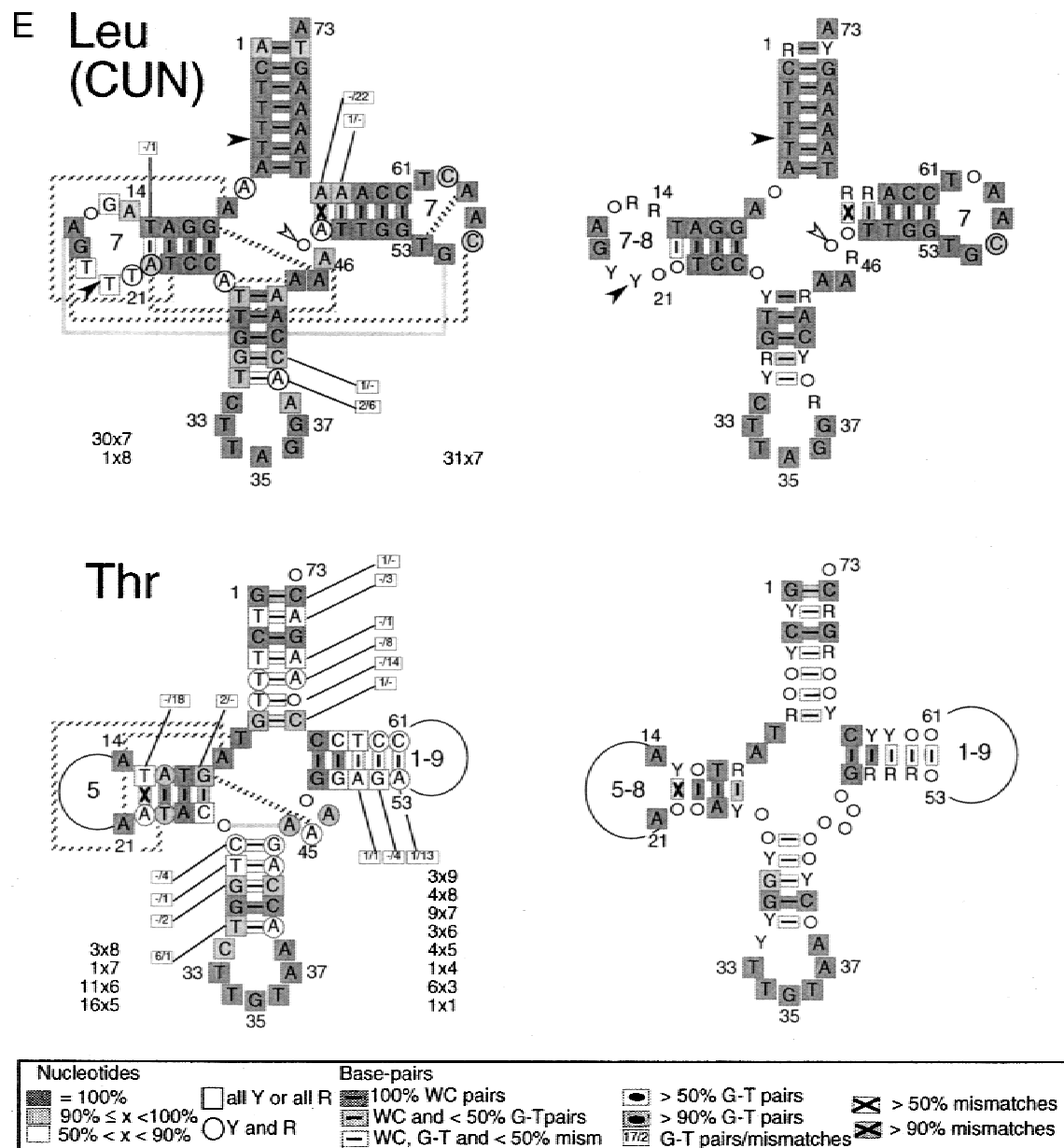


FIGURE 2. (Continued.) Mammalian mt tRNAs as deduced from sequence comparison within 31 mammalian mitochondrial genomes. Sequence features of each tRNA family are summarized in two cloverleaves, on the left typical tRNAs, and on the right consensus tRNAs. tRNAs have been arranged according to their order of appearance in the DNA strands: "light" tRNAs are those transcribed from the heavy DNA strand and have a low G content, whereas "heavy" tRNAs are those transcribed from the light strand DNA and thus have a high G content (Anderson et al., 1981). Amino acid specificities of tRNAs are according to their anticodon sequence. Numbering of nucleotides is according to Sprinzl et al. (1998) and takes into account gaps in D- and T-loops due to length variations. In the left panel, nucleotides present at a given position in more than 50% of the tRNAs are indicated, those present at lower levels are represented by an open circle or a line. Nucleotides fully conserved have a dark gray background, those present at more than 90% are on light gray background, and those present at more than 50% but less than 90% are on a white background. Conservation within the same family of either purines or pyrimidines is indicated by the squared shape of the background, whereas nonconservation in such a family is indicated by the circled shape of the background. Secondary interactions are indicated by a straight line for classical Watson-Crick base pairs, an elongated black dot for G-T base pairs, and a cross for mismatches. These signs are indicated whenever more than 50% of the cases are concerned. For classical Watson-Crick pairs, dark gray backgrounds correspond to a 100% conservation, white backgrounds to the presence of at least 1 mismatch, and light gray to the presence of at least 1 G-T pair. Whenever G-T pairs or mismatches are present at 50% or more, they are indicated by their specific signs with a dark gray background when present at 100%. The number of G-T pairs and mismatches at a given base pair are indicated in detail for each position by two numbers in an opened square, the number on the left referring to G-T pairs and the number on the right referring to mismatches in these domains. Dark arrowheads indicate positions where 1 additional nt is found in some sequences, and open arrowheads indicate positions where 1 nt is missing. The large variability in size of the D- and T-loops did not allow us to easily define conserved nucleotides. The average length of the loops is indicated within the loops, and a detailed analysis of the variability of the loops is given below each loop. (Figure continues on next page and legend continues on page 1367.)

F “Heavy” tRNAs

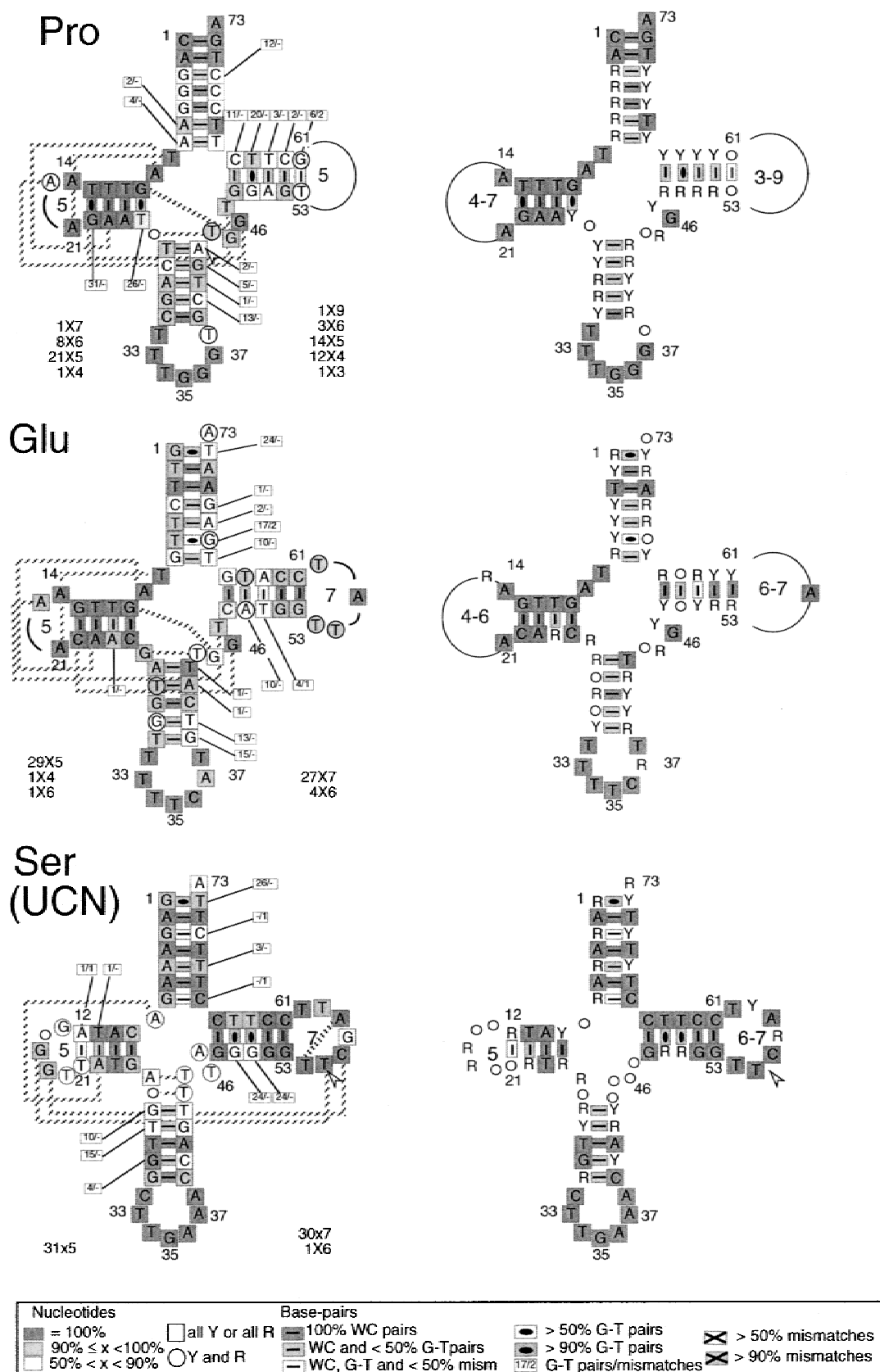
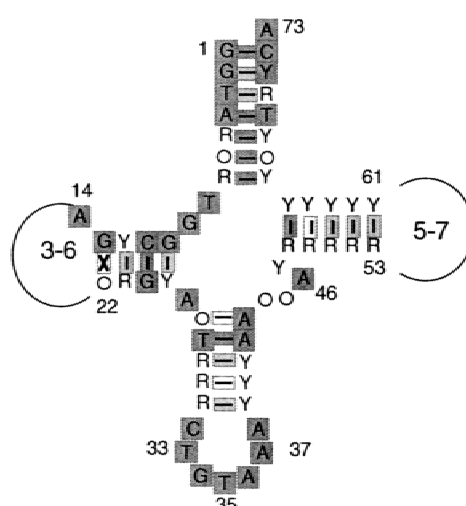
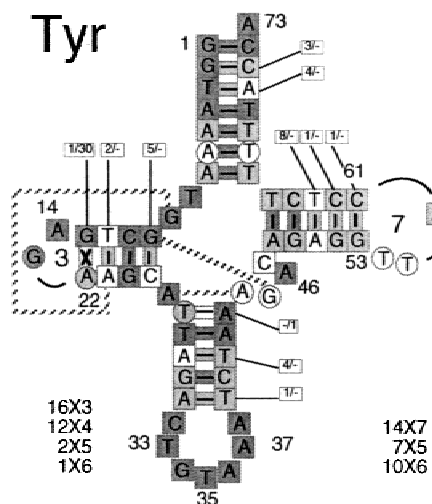
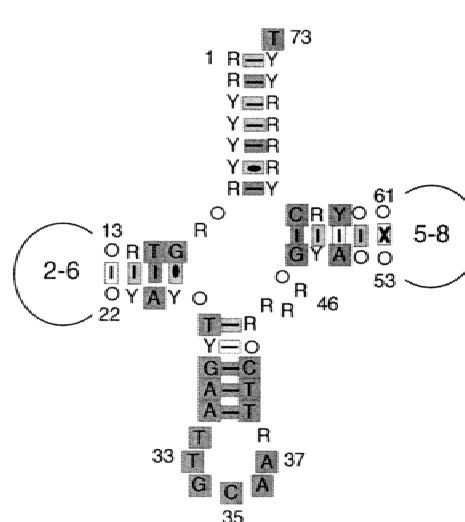
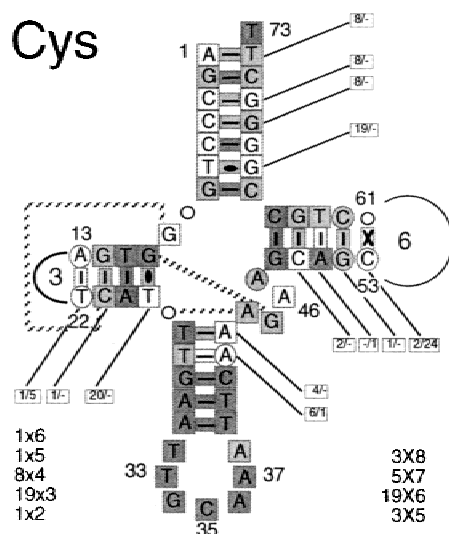


FIGURE 2. (Continued.)

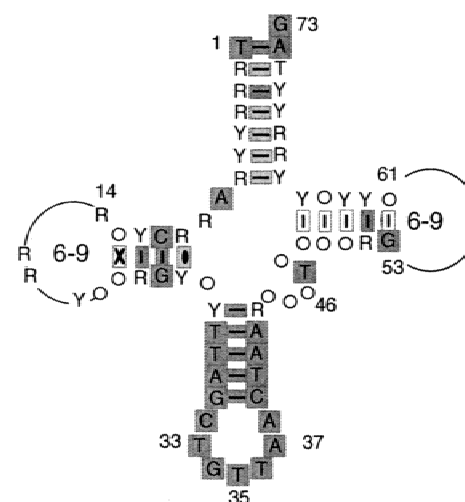
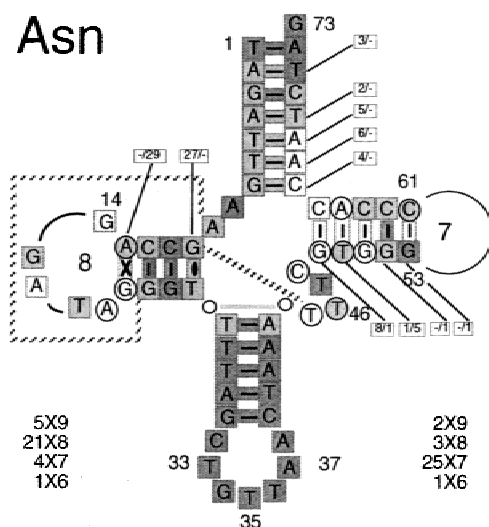
G Tyr



Cys



Asn



Nucleotides		Base-pairs	
= 100%	all Y or all R	100% WC pairs	> 50% G-T pairs
90% ≤ x < 100%	Y and R	WC and < 50% G-T pairs	> 90% G-T pairs
50% < x < 90%		WC, G-T and < 50% mism	G-T pairs/mismatches
			> 50% mismatches
			> 90% mismatches

FIGURE 2. (Continued.)

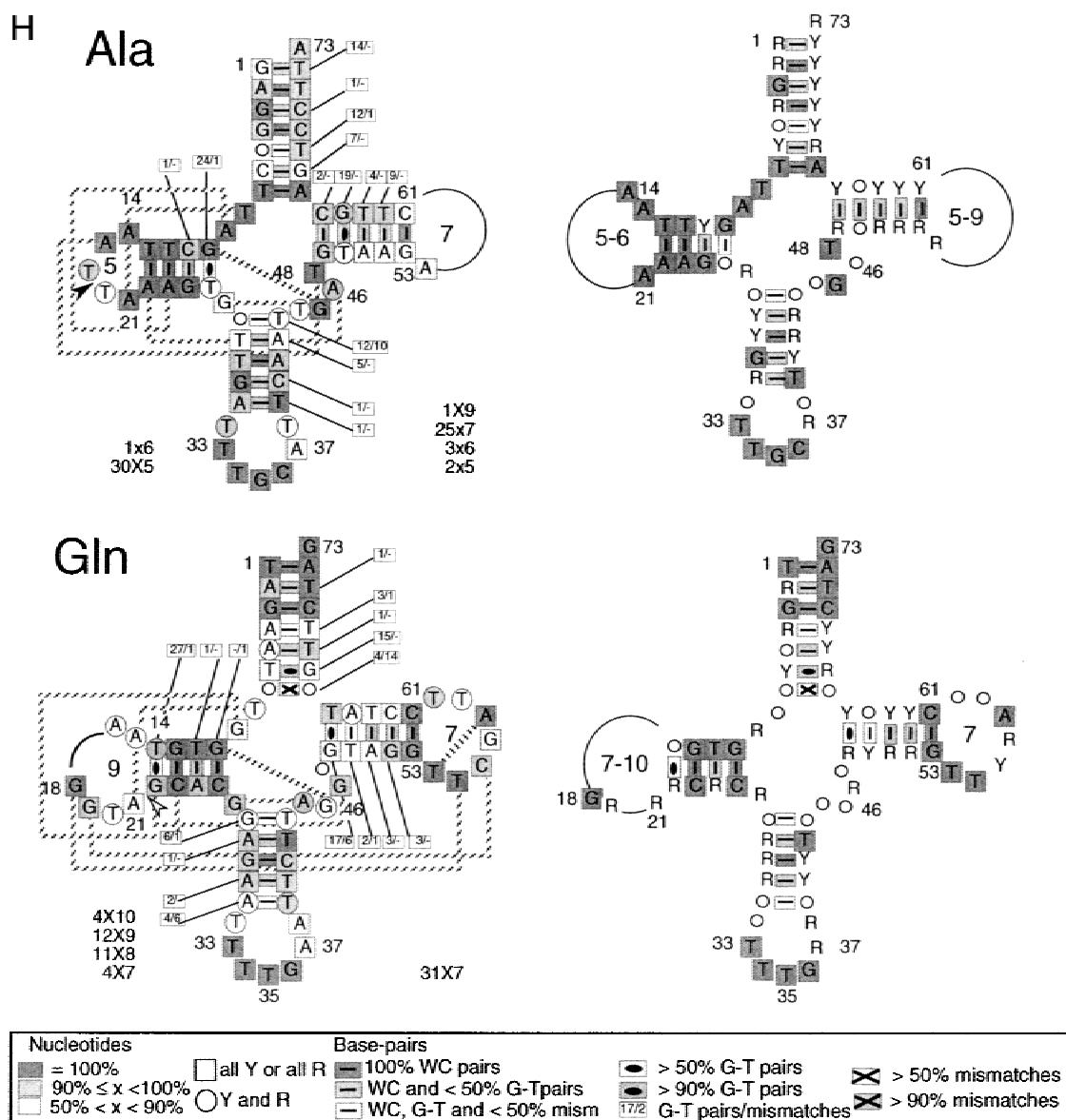


FIGURE 2. (Continued.) Potential tertiary interactions are indicated by stippled lines, by analogy to classical tRNA interactions as well as to those established for mitochondrial tRNA^{Phe} (Wakita et al., 1994). In the right panel, "consensus" mammalian mt tRNAs are displayed. Fully conserved nucleotides are indicated on dark gray squared backgrounds. Conservation of purines is indicated as R and conservation of pyrimidines is indicated as Y. Nonconserved nucleotides are indicated by open dots. Conservation of secondary interactions are as in the left panel. Minimal and maximal lengths of D- and T-loops are recalled.

(Giegé et al., 1993) tRNA^{Leu}, tRNA^{Ser}, and prokaryotic tRNA^{Tyr}, are known to have large variable loops (up to 23 nt).

Size of the stems

Canonical tRNAs have D-stems varying in size from 3 to 4 bp. Most of the mammalian tRNAs analyzed here can be drawn with a 4-bp D-stem, with a strong conservation of size within a given tRNA family. In five families (Leu(UUR), Trp, Gly, Pro, and Gln), the base

pair located next to the loop (nt 13 and 22 according to conventional tRNA numbering) is typically G-T and in four other families (Ile, Thr, Tyr, and Asn), it is typically a mismatch. Although G-T pairs would actually be formed at these positions and be stabilized by the neighboring triple interaction T8-A14-A21 (see below), mismatches may likely lead to shortening of the D-stems and enlargements of the loops. This is especially likely in tRNA^{Tyr}, where the D-loops have sizes too small to overcome stereochemical constraints to allow for a loop to form. Also, in some tRNA families where nt 13 and

22 can base pair, however, pairing may not occur because the D-loop is formed only by a very small (three or less) number of nucleotides (all tRNA^{Lys}, most tRNA^{Cys}, three species of tRNA^{Asp} and three species of tRNA^{Tyr}). Interestingly, mismatches formed by purine–purine pairs between nucleotides 13 and 22 are typical for class II canonical tRNAs (Giegé et al., 1993) and might thus be a leftover of the bacterial ancestor tRNA in the case of mt tRNA^{Tyr}.

T-stems are typically defined by 5 bp in canonical tRNAs. As seen in Figure 2, this situation is also found in a number of mt tRNAs. However, bp 61–53, next to the T-loop, is in some instances formed by either a G–T pair or corresponds to a mismatch. This is typically the case for the tRNA^{Cys} family with a conserved mismatch, but also in 14 tRNA^{Phe} and tRNA^{Thr} sequences, 10 tRNA^{Trp}, 8 tRNA^{Pro}, 8 tRNA^{Val}, 7 tRNA^{His}, 6 tRNA^{Ser(AGY)}, and 1–4 sequences of the Asp, Lys, Gly, Arg, Tyr, and Asn families. Here again, a smaller T-stem formed only of 4 bp and an enlarged T-loop are likely to occur. Brulé et al. (1998) have discussed the opening of a terminal G53–U61 base pair in human mt tRNA^{Pro}, in favor of a seven-membered classical T-loop that shows an appropriate posttranscriptional modification pattern. An opening of the T-stem would be especially favorable for those tRNAs with 1–3-nt-long T-loops (Phe, Val, Thr, Pro, Arg).

Typical mismatches present at strategic positions, that is, at the end of stems, are noteworthy. Thus, in a large number of tRNA^{Val} (18 out of 31) and tRNA^{Leu(CUN)} (22 out of 31), nt 49–65 (at the other end of the T-stem) form a mismatch. This likely does not interfere with the length of the stem, but rather introduces an irregularity in the stacking of the acceptor stem on the T-stem in the three-dimensional folding of those tRNAs. In some tRNA^{Leu(UUR)}, a mismatch between nt 31 and 39 at the bottom of the anticodon stem may shorten the stem and enlarge the loop.

Size of the D- and T-loops

Variability of D-loop size is well known in classical tRNAs. However, whereas a typical D-loop size in classical tRNAs is 8–10 nt (Sprinzl et al., 1998), the average and most frequently encountered D-loop size in mammalian mt tRNAs is 5 nt (Fig. 3A). Noteworthy here in the case of mammalian mt tRNAs is not only the large variability in size from one tRNA family to another, but also the large variability within a given family (Fig. 2). Extreme situations are found in the Phe, Gly, Trp, Val, and Cys families, with sizes ranging from 5 to 11, 5 to 10, 5 to 9, 5 to 9, and 2 to 6 nt, respectively. Alternatively, there are also families where the size does not vary or varies only weakly (Ile, Ser(UCN), Lys, Ala, His, Met, Leu(UUR), and Leu(CUN)). As will be discussed later, invariant large D-loops in Leu(UUR), Leu(CUN), and to a certain ex-

tent Gln correlate with tertiary interactions between this loop and the T-loop.

Conservation of T-loop size in canonical tRNAs (7 nt) is a well-established fact (Sprinzl et al., 1998). This is not the case for many mt tRNAs. Although the seven-membered loops form a large majority, T-loops as small as 1–2 nt and as large as 10 nt are found (Fig. 3A). As is the case for D-loops, variability in T-loop sizes is taking place within specific families as well as between the different families.

It must be noted that there is no apparent correlation between D- and T-loop sizes within tRNAs of the same family, neither at the level of the mean size nor at that of the variability in size.

Stability of secondary interactions

As seen in Figure 2 and summarized in Table 2, certain secondary interactions in each of the 22 tRNA families studied here tolerate either G–T (or T–G) base pairs or mismatches in parallel to classical Watson–Crick pairs. Within a total of 14,197 base pairs analyzed, 908 (6.4%) correspond to G–T pairs and 717 (5%) to mismatches (Table 2). “Light” tRNAs, that is, those transcribed from the heavy-strand mitochondrial DNA, are globally poor in G–T (2.1%) and rich in mismatches (6.2%), whereas the “heavy” tRNAs are G–T rich (13.6%) and poor in mismatches (3%). This analysis also takes into account those G–T pairs and mismatches localized at the end of the D- and T-stems, which actually may be disrupted in the secondary structures. This is typically the case for tRNA^{Leu(UUR)}, tRNA^{Trp}, and tRNA^{Gly}, which are richer in G–T pairs than the average light tRNAs or for tRNA^{Tyr}, tRNA^{Cys}, and tRNA^{Asn}, which typically must have an opened mismatch at the end of these stems because of the very short sizes of the neighboring loops. Thus, the ratio between G–T and mismatches in each family might even be more pronounced. This distribution is typically linked to the nucleotide composition of each of the mitochondrial DNA strands [about 32% A, 30% C, 13% G, and 25% T for the light strand in mammalian mt DNA (Anderson et al., 1981; Boore, 1999, and references therein)]. Although the nucleotide composition of tRNA genes deviates somehow from these values (analysis of three heavy tRNAs, i.e., Ile, His, and Phe, leads to 12.9–17.3% G residues and analysis of three light tRNAs, i.e., Gln, Cys, and Tyr leads to 12.4–19.4% C residues), the G-richness or G-poverty is maintained and fits with the fact that G-rich sequences allow for higher numbers of G–T base pairs than G-poor sequences.

Analysis of the distribution of G–T pairs and mismatches within the four domains of the complete set of analyzed tRNAs reveals further interesting features. Figure 3B summarizes the distribution along the two main domains of the potential L-shaped three-dimensional structure of the tRNAs, namely the accep-

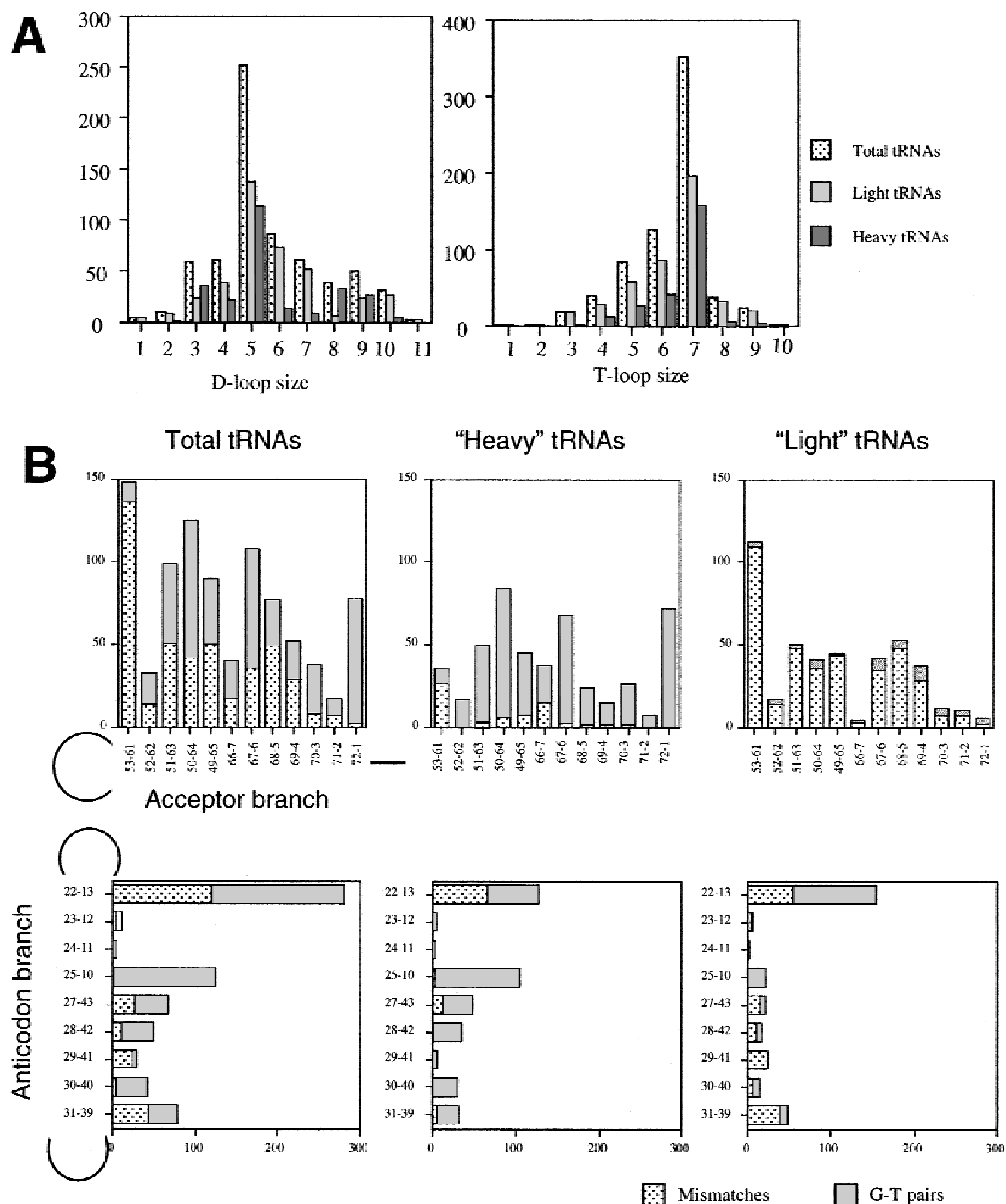


FIGURE 3. A: Statistical distribution of D- and T-loop sizes in 648 sequences (all except tRNA^{Ser(AGY)}). **B:** Number and statistical distribution of G-T base pairs and mismatches along the acceptor branch and the anticodon branch of the same tRNAs. The two panels on the left summarize the situation in all tRNAs, the two central panels describe the case of the heavy tRNAs (401 sequences) and the two panels on the right describe the case of the light tRNAs (247 sequences). The gray part of the histograms highlights mismatches and the dotted parts correspond to G-T pairs. The two branches of the tRNAs are represented by the names of their base pairs. Thus, the acceptor branch contains 12 bp numbered from bp 53–61 to bp 1–72 neighboring the T-loop and the CCA terminus, respectively. The anticodon branch is formed of 9 bp numbered 13–22 to 31–39 and ranging from the D-loop down to the anticodon loop (see Fig. 1).

TABLE 2. Base pairing within the different mammalian mitochondrial tRNA families.

	Total base pairs	Watson–Crick	G–T	Mismatches
Light tRNAs				
Phe	651	570	6	75
Val	651	590	8	53
Leu(UUR)	651	585	33	33
Ile	651	612	1	38
Met	651	559	0	92
Trp	651	585	33	33
Asp	651	617	18	16
Lys	609	585	12	12
Gly	651	603	38	10
Arg	651	621	11	19
His	651	619	13	19
Ser(AGY)	558	494	5	59
Leu(CUN)	651	618	4	29
Thr	651	569	12	70
Total	8979	8227 (91.6%)	194 (2.1%)	558 (6.2%)
Heavy tRNAs				
Pro	651	511	138	2
Glu	651	549	99	3
Ser(UCN)	682	571	108	3
Tyr	651	590	30	31
Cys	630	520	80	30
Asn	651	557	57	37
Ala	651	517	112	22
Gln	651	530	90	31
Total	5218	4345 (83.3)	714 (13.6%)	159 (3.0%)

This analysis refers to the 7 bp from the acceptor stems, 4 bp of the D-stems, 5 bp of the anticodon stems (6 for tRNA^{Ser(UCN)}), and 5 bp in the T-stems. Possible enlargement of loops has not been taken into account in these statistics. Light tRNAs are those transcribed from the heavy strand mitochondrial DNA and Heavy tRNAs are those transcribed from the light strand.

tor branch (stacking of the acceptor stem and the T-stem) and along the anticodon branch (stacking of the D-stem on the anticodon stem). When considering the whole set of tRNAs (total tRNAs), G–T and mismatches appear at any position, except at the level of bp 12–23 and 11–24 in the D-stem. The almost perfect conservation of stable Watson–Crick pairs at these two positions can be considered as the most conserved secondary structure interactions of mammalian mt tRNAs and suggests their great importance in the architecture of these tRNAs. In contrast, bp 53–61 and 13–22 are the positions most often occupied by noncanonical interactions, which reflects again the possibility of shortening of the stems and enlargement of the D- and T-loops in the corresponding tRNAs. Distribution of the tRNA according to their genomic origin illustrates the G–T richness of the heavy tRNAs, and the richness in mismatches of the light tRNAs. Interestingly, whereas in light tRNAs, bp 7–66 and 1–72 are those with the least deviations from Watson–Crick pairs, the corresponding pairs in the heavy tRNAs, and especially bp

1–72, are much more prone to G–T pairs. In the anticodon branch, the main difference between light and heavy tRNAs is at positions 10–25, where G–T pairs are often seen in heavy tRNAs. G10–T25 base pairs are not unusual in classical tRNAs either, and have been structurally characterized by crystallography of yeast tRNA^{Asp}, where they were shown to participate in base triples involving nucleotides in the variable loop (Westhof et al., 1985). The G10–U25 in this tRNA was shown to be important for aminoacylation identity, probably for structural reasons (Pütz et al., 1991).

Finally, detailed analysis of the location of mismatches in the individual 22 tRNA families (Fig. 2) reveals that, although these non-Watson–Crick interactions are rather spread out in the secondary structures of most tRNA families, they are highly concentrated to unique positions in the middle of domains in others, including Phe, Met, Ser(AGY), Leu(CUN), and Gln. Most of the spread out mismatches are of a random nature, suggesting that they do not fulfill any specific function, but, because of the absence of selective pressure, they are allowed to accumulate. In contrast, those restricted to selected positions, as for example those present in the acceptor domain of tRNA^{Met}, are highly conserved also at the level of the primary sequence and must have been retained by evolution and be of structural and/or functional importance. The same holds true for the G–T base pairs. To be noted are the two G–T pairs in the T-stem of tRNA^{Ser(UCN)}, which may form a specific signal, selected and conserved during evolution.

Search for tertiary features

The three-dimensional typical L-shaped structure of classical tRNAs is based on a set of nine long-range tertiary interactions, mainly between conserved or semiconserved nucleotides, and on the presence of conserved nucleotides at strategic positions, leading to local structural information (Fig. 1). Anticodon loop nucleotides U33, Y32, and R37 and base pair G53–C61 at the end of the T-stem belong to the latter nucleotides. Thus, U33 confers a turn to the phosphodiester backbone, the so-called “U-turn,” and induces a typical structure in the anticodon loop. Long-range interactions take place between residues U8–A14–A21, 9–23–12, 25–10–45, 13–22–46, R15–Y48, G18–Y55, G19–C56, 26–44, and T54–A58 (Fig. 1).

tRNA^{Ser(AGY)} and tRNA^{Ser(UCN)}

As mentioned above, both serine isoacceptor tRNAs from mammalian mitochondria have unusual secondary structures. The present analysis shows that this is the case in the 31 mammalian genes compared here. Thus, all tRNA^{Ser(AGY)} lack the complete D-arm and all tRNA^{Ser(UCN)} are missing a single nucleotide in the connector between the acceptor and the D-arm. The three-

dimensional structures of tRNA^{Ser(AGY)} have been investigated in the case of bovine and human (Ueda et al., 1983; Hayashi et al., 1997b), and that of tRNA^{Ser(UCN)} in the case of bovine (Watanabe et al., 1994a; Hayashi et al., 1998). According to our compilation, all mammalian tRNA^{Ser(AGY)} on one hand, and all tRNA^{Ser(UCN)} on the other hand, do fold into similar three-dimensional structures. From the fact that they function in the same protein-synthesizing environment, one can infer that they possess similar overall structures. Especially interesting in this regard is their recognition by a single mt seryl-tRNA synthetase (Ueda et al., 1992; Hayashi et al., 1998).

Tertiary features in "classical" mitochondrial tRNAs

The presence of structurally important nucleotides and the possibility of forming "classical" tertiary interactions have been investigated in the 20 mammalian mt tRNA families sharing a classical cloverleaf. Interestingly, residue T33 in the anticodon loop is fully conserved in all tRNA families, with the noticeable exception of two valine-specific tRNAs and the complete methionine family, which presents a strictly conserved C33. Interestingly, the few classical tRNAs with C33 instead of U33 are initiator tRNAs (Sprinzl et al., 1998). Also valylatable viral tRNA-like structures, not involved in protein synthesis, typically lack a U33 (Florentz & Giegé, 1995).

In all 20 families, R37 and Y32 are most often present. In summary, the great majority of tRNAs do have an anticodon loop with a classical structure, with the noticeable exception of the tRNAs specific for methionine.

Figure 2 and Table 3 highlight potential tertiary interactions in each family. Interestingly, these interactions are rather numerous in most families. The triples 25–10–45 and 9–23–12 as well as interaction 26–44 can be formed everywhere with combinations of primary sequences found in canonical tRNAs (Giegé et al., 1993). Also, the triple T8–A14–A21 is often found, with only one clear exception (tRNA^{Cys}) and two likely exceptions, namely tRNA^{Lys} and tRNA^{Tyr}, where the D-loop may be too small and may not allow for these interactions. The last type of triple interaction, 13–22–46, is possible in 14 families but is likely to be excluded in those tRNAs where there is a mismatch between residues 13 and 22 (see above).

Other long-range interactions involving specific nucleotides from the D- and T-loops (G18–Y55, G19–C56, and T54–A58) are typically not present in mammalian mt tRNAs. Only in tRNA^{Leu(UUR)} and tRNA^{Gln} do these three interactions exist. For tRNA^{Leu(CUN)}, both loops are remarkably well conserved in size and rather well conserved in sequence. The T-loop has conserved T54, G55, C56, and A58 (only *Myoxus glis* has a T54GAACT60 T-loop with A56 instead of C56), whereas the D-loop has fully conserved A18 and G19.

TABLE 3. Potential tertiary interactions in typical mammalian mitochondrial tRNAs.

tRNA	18–55	19–56	(8–14)–21	9–(23–12)	15–48	(25–10)–45	(13–22)–46	26–44	54–58
Light tRNAs									
Phe	—	—	TA-A	A-AT	AT	CG-G	TA-A	AA	—
Val	—	—	TA-A	A-AT	AT	CG-G	TA-A	AA	—
Leu(UUR)	GT	GC	TA-A	G-TA	GC	CG-A	GT-N	AT	TA
Ile	—	—	TA-A	G-AT	—	AT-T	mism	GA	TA
Met	—	—	TA-A	C-GC	AT	TA-A	TA-A	AA	<i>T54A57</i>
Trp	—	—	TA-A	A-AT	AC	CG-A	TG-G	AA	—
Asp	—	—	TA-A	A-TA	AT	CG-A	AT-T	AA	—
Lys	—	—	<i>AA-T</i>	A-AT	—	CG-G	TA-A	AA	—
Gly	—	—	TN-A	A-TA	—	CG-A	TG-G	AT	—
Arg	—	—	TA-A	A-AT	AT	CG-G	TA-A	AA	—
His	—	—	TA-A	A-AT	AT	CG-A	TA-A	AN	—
Ser(AGY)	—	—	—	—	—	—	—	—	—
Leu(CUN)	AG	GC	<i>AA-T</i>	A-TA	—	CG-A	TA-A	AA	TA
Thr	—	—	TA-A	A-TA	—	CG-A	mism	NA	—
Heavy tRNAs									
Pro	—	—	TA-A	A-AT	AT	TG-G	TG-G	NA	—
Glu	—	—	TA-A	A-AT	AT	CG-G	GC-G	GT	<i>T54A57</i>
Ser(UCN)	GT	GC	—	A-AT	—	GC-T	—	GT	TA
Tyr	—	—	TA	G-AT	—	CG-G	mism	AA	<i>T54A57</i>
Cys	—	—	—	G-CG	—	TG-G	—	NA	—
Asn	<i>GN</i>	<i>AN</i>	<i>AG-A</i>	A-GC	<i>TC</i>	TG-T	mism	NN	—
Ala	—	—	TA-A	A-AT	AT	TG-G	TA-A	GT	—
Gln	GT	GC	TA-A	G-CG	<i>AN</i>	CG-G	TG-G	GA	TA

Tertiary interactions present in canonical tRNAs have been sought. Sequences in italic highlight nonclassical cases, that is, those not found in canonical tRNAs (Giegé et al., 1993). *N* stands for any nucleotide and mism for the presence of a mismatch instead of a base pair.

Thus interactions T54–A58 and G19–C56 can take place. The strict covariation between A18 and G55 indicates a new type of tertiary D-loop/T-loop interaction in this tRNA family. A related type of D-loop-T-loop interaction might be present in mt tRNA^{Asn}. The D-loop is much larger than the average five-membered one, as in the leucine and glutamine tRNAs, and there is also a predominant, even though not strictly conserved purine–purine G18A19 motif. A T-loop of 7 nt is clearly predominant, but not conserved as in the other cases. In tRNA^{Ile}, only the interaction T54–A58 is present. In three additional tRNA families (Met, Glu, and Tyr), this interaction could be substituted by a T54–A57 interaction.

The search for alternative D-loop/T-loop interactions in the remaining tRNA families was rather difficult because of the large variability in size and sequences of these loops. A statistical analysis of nucleotide composition of both loops (Table 4) did not find typical covariations. However, it did highlight typical features for both light and heavy tRNAs. Indeed, all light tRNAs have a large deficit in G residues both in their D- and T-loops (with the exception of both tRNA^{Leu}). The percentage of G is at most 5.7%, and thus much lower than the global G content of mt DNA light strands (~13%) and than tRNA genes (12%–17%). Thus, the poverty in G of both loops is a typical signature of light tRNAs.

Similarly, heavy tRNAs have D-loops that are very poor in C residues (<5.8%).

Summarizing the analysis of potential tertiary features, it appears that only three or four mammalian mt tRNA families have the possibility of folding into a three-dimensional structure according to the same structural rules as classical tRNAs. These are the Leu(UUR), Leu(CUN), Gln, and possibly Asn families. The tRNA^{Ser(UCN)} family has similar rules. For all other families (with the exception of Ser(AGY)), classical tertiary interactions are possible only within the core of the tRNAs, excluding D- and T-loop interactions. In seven tRNA families (Trp, Gly, Thr, Ile, Tyr, Lys, and Cys), it was only possible to find tertiary interactions within the central part of the tRNA cloverleaves. Thus, the corresponding tRNAs either have weaker tertiary structures or are folded on the basis of alternative interactions. Figure 4 summarizes the various situations and highlights the presence of a minimal set of three tertiary interactions present in all tRNA families, namely between nt 23–12–46, 25–10–45, and 26–44, as well as the presence of a strongly conserved Watson–Crick secondary interaction between nt 24 and 11. The stacking of these four sets of base combinations likely represents the minimal level of tertiary information required for functional activity of the mammalian mt tRNAs.

TABLE 4. Statistical analysis of D- and T-loops in mammalian mitochondrial tRNAs.

tRNA	Loop size		Nucleotide content (%)							
	D-loop	T-loop	D-loop				T-loop			
	(average and limits)		A	C	G	T	A	C	G	T
Light tRNAs										
Phe	9 (5–11)	5 (2–7)	51.1	26.4	1.6	21.9	32.3	32.3	5.9	29.4
Val	5–6 (5–9)	5 (3–6)	64.8	18.1	1.7	15.4	43.5	25.1	5	26.4
Leu(UUR)	10 (9–10)	7	34.6	25	<u>30.4</u>	9.9	34.5	20.3	0.5	44.7
Ile	4	7 (6–8)	74.2	13.7	0.8	11.3	40.1	18.9	3.7	37.3
Met	5 (5–6)	7 (6–7)	75.1	5.7	0.6	18.5	32	16.3	0	51.7
Trp	5–6 (5–9)	5 (4–6)	62.5	24.6	2.1	10.7	52	18.9	4.1	25
Asp	5 (3–6)	7 (5–7)	71.3	17.8	0.8	11.6	44.2	16.5	7.5	29.6
Lys	3 (1–3)	7 (4–10)	44.3	20	5.7	30	39.5	30.2	5.1	25.1
Gly	5–6 (5–10)	6 (6–7)	59	14.2	3.9	22.7	44.3	21.8	3.6	30.2
Arg	6 (5–7)	6 (3–7)	70.5	19.8	2.8	6.8	37.9	16.8	3	42
His	5 (4–5)	7 (5–9)	76	17.6	1.4	4.5	42	28	4.1	25
Ser(AGY)	—	8 (7–9)	—	—	—	—	50.6	22.4	3.8	23.1
Leu(CUN)	7 (5–8)	7	42.3	10	<u>25.7</u>	22	29.5	27.1	<u>14.3</u>	29
Thr	7 (5–8)	7 (1–9)	66.6	11.3	3.9	18	39.3	33.8	5.5	21.3
Heavy tRNAs										
Pro	5 (4–7)	5 (3–9)	61.5	0.6	9.7	28	32.4	2.7	26.4	38.5
Glu	5 (4–6)	7 (6–7)	70.9	0	5.2	23.9	28.1	1.4	19.2	51.2
Ser(UCN)	5	7 (6–7)	16.8	5.8	<u>54</u>	42.5	15.7	15.7	12.5	56
Tyr	3 (3–6)	7 (5–7)	41	0.9	<u>30.3</u>	27.6	33.5	10.5	27.3	28.6
Cys	3 (2–6)	6 (5–8)	38.2	1.9	6.9	52.9	35.7	19.2	7.1	35.7
Asn	8 (6–9)	7 (6–9)	25.1	4.8	25.1	44.9	32.3	2.2	30	35.4
Ala	5 (5–6)	7 (5–9)	62.8	0	1.9	35.2	34.9	1.4	32.5	34.9
Gln	8–9 (7–10)	7	56.5	2.2	10.3	30.9	21.6	15.2	10.5	52.3

This analysis compiles 679 tRNA sequences. Striking values are given in bold, exceptions to these are underlined.

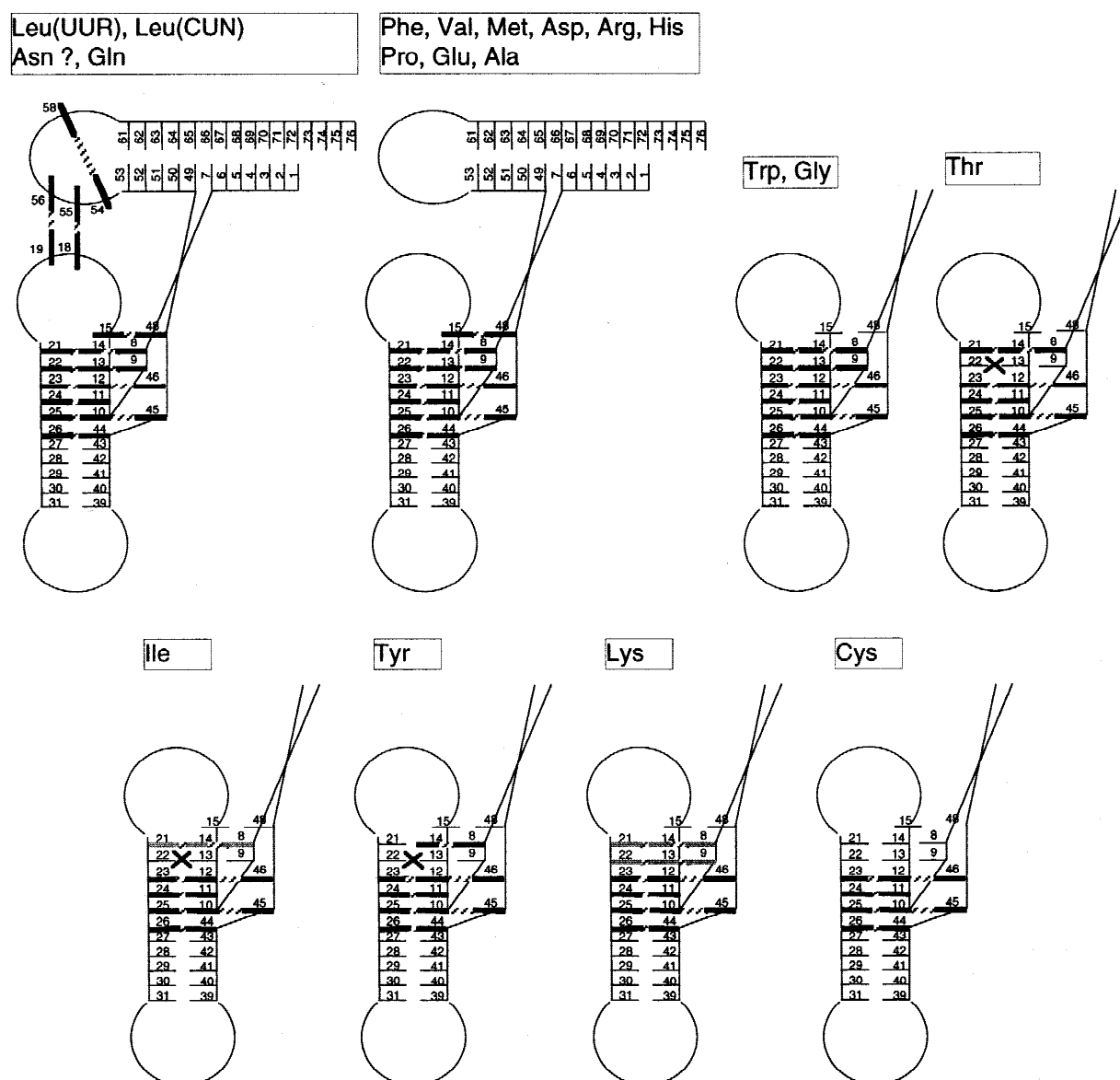


FIGURE 4. Potential tertiary interactions in mammalian mt tRNA families, highlighted on three-dimensional L-shaped structures. The case of the two tRNA^{Ser} families, whose structures have been established experimentally (see text for references), are not shown. Families have been classified according to the number of tertiary interactions potentially present. Nucleotides are represented by lines and the corresponding numbering in canonical tRNAs (Sprinzl et al., 1998). Nucleotides involved in tertiary interactions are in bold lines. Only four families (Leu(UUR), Leu(CUN), Gln, and Asn) have the full set of the nine interactions found in canonical tRNAs. All other families are missing some of the nine interactions and, in particular, those between nucleotides in the D- and T-loops. Thus, only the anticodon branch of the L is shown for these families.

Conservation of primary structure

Mammalian mitochondrial tRNA families include a large range of primary sequence variations

Although the cytosolic tRNAs among mammals are almost completely identical [e.g., tRNA^{Lys3} is identical down to the modified nucleotides among human, bovine, rabbit, and rat (G. Keith, pers. comm.)], no pair of identical mt tRNAs could be found in our set of 679 mt

tRNA genes. This illustrates that primary sequences are not well conserved. However, the present vertical alignments detected many more conserved elements than did the previous horizontal alignments, and indeed, in each tRNA family (Fig. 2), a number of strictly conserved nucleotides can be found. Strong conservation is found not only at individual positions, but also over large structural domains in some tRNAs.

The most striking case concerns the methionine tRNAs, which have the largest number of nucleotides

fully conserved over the 31 mammalian species. Also, the isoleucine and the two leucine families have a large number of strictly conserved nucleotides all over their structures. Most tRNA families show high conservation within individual domains of the secondary structure. Thus, often, the D-stem nucleotides, anticodon loop, and anticodon stem nucleotides are fully conserved, or alternatively, the T-stem nucleotides are strictly conserved. At the other end of the scale, aspartic acid-specific tRNAs share the lowest number of conserved nucleotides.

A less stringent analysis (Fig. 2, right panels) considering not only fully conserved nucleotides but also semiconserved elements (i.e., purines or pyrimidines) shows relatively little deviation from one tRNA to another inside each family. At most, 35% of the individual positions (excluding the D- and T-loops) are neither conserved nor semiconserved in the worst cases (Phe, Val, and Thr families). This level falls below 10% for the Ile, Leu(CUN), Pro, and Tyr families and to 1.7% for the Met family.

The specific case of tRNA^{Met}

The outstanding level of primary sequence conservation within the methionine tRNAs, and as a consequence, the high conservation in secondary and tertiary structures of these tRNAs, make the tRNA^{Met} family a unique case. Interestingly, there is only one type of tRNA^{Met} encoded in the mammalian mt DNAs, so that a single molecule has to serve both as initiator and elongator tRNA. Thus, apart from being methionylated by methionyl-tRNA synthetase, this single tRNA must on one hand, be recognized by translation initiation factors and the mt methionyl-tRNA transformylase, and be able to read start codons of the mt mRNAs, and, on the other hand, the unformylated methionyl-tRNA must be specifically recognized by the elongation complex and be able to decode elongation codons. This need for a relatively high number of different interactions undergone by mt tRNA^{Met}, as compared to the other mt tRNAs, is likely the reason for their high degree of conservation. Each interaction with a specific protein would require different conserved features. The two conserved T50–T64 and Y51–N63 mismatches in the T-stem and a A5–R68 mismatch in the acceptor stem are features that have so far not been identified in any other tRNA, and are therefore prime candidates as signals for recognition (or against recognition) by either of the above mentioned proteins.

Among the different deviations from the universal genetic code found in mitochondria (Watanabe & Osawa, 1995) is the AUA codon, otherwise assigned to isoleucine, which has been taken over by methionine. A unique posttranscriptional modification of C34 has been identified as 5-formylcytidine and is thought to be responsible for the unique decoding properties (Moriya et al.,

1994; Takemoto et al., 1995). Because the mammalian initiation and elongation factors (Ma & Spremulli, 1995; Zhang & Spremulli, 1998) and the mt methionyl-tRNA transformylase (Takeuchi et al., 1998) are known, the MetRS seems to be the only missing piece to solve these highly interesting mechanisms.

General functional implications

Conserved or semiconserved nucleotides between tRNAs of a same specificity but within different mammalian species, may not only be involved in the establishment of structural characteristics (as discussed above), but also in specific recognition by molecular partners. Whereas all tRNAs have to interact in a non-specific way with a large series of proteins, including maturation enzymes (i.e., RNaseP and tRNA-nucleotidyltransferase), some posttranscriptional modification enzymes, elongation factor, and ribosomal proteins, each is also specifically recognized by several enzymes. These include, of course, the cognate aminoacyl-tRNA synthetase that catalyzes the precise charging with the cognate amino acid, but also very specific posttranscriptional modification enzymes.

Not much is presently known about posttranscriptional modifications occurring in mammalian mt tRNAs, or about signals within canonical tRNAs, necessary for specific recognition by the corresponding enzymes (Grosjean & Benne, 1998). In contrast, aminoacylation identity elements (i.e., the set of nucleotides and structural features within tRNAs) required for specific recognition and aminoacylation by the cognate aminoacyl-tRNA synthetase are well defined for a large series of classical tRNAs (reviewed in McClain, 1995; Giegé et al., 1998). Identity sets for all 20 *Escherichia coli* specificities as well as for a number of other prokaryotic and even some mitochondrial systems are established. They are mainly composed of a limited number of nucleotides concentrated in the anticodon domain and the end of the acceptor stem, next to the CCA 3' terminus. These sets are rather conserved in evolution (McClain, 1995; Giegé et al., 1998).

Aminoacylation identity elements for mammalian mt tRNAs are only known for tRNA^{Asp/Gly} (Börner et al., 1996) and tRNA^{Ser} (Ueda et al., 1992). On the basis of (1) the endosymbiotic hypothesis of the origin of mitochondria, (2) conservation of major identity elements over various kingdoms, and (3) the ability of mt aminoacyl-tRNA synthetases to aminoacylate *E. coli* tRNAs (Kumazawa et al., 1991), it was tempting to search for potential identity elements within the different tRNA families analyzed here. Table 5 recalls the *E. coli* identity sets (according to Giegé et al., 1998) and lists the corresponding elements present in the typical tRNA sequences for the 20 different aminoacylation specificities in mammalian mitochondria. Interestingly, in two systems (Tyr and Asn), all elements of

TABLE 5. Search for the presence of *E. coli* tRNA aminoacylation identity elements in mammalian mitochondrial tRNAs.

tRNA	Acceptor stem	Anticodon loop	Other domains
Phe	A73	G34, A35, A36	G27:C43, G28:C42, U20, G44, U45, U59, U60
Val	A73 , G3:C70, U4:A69	A35, C36	—
Leu	A73	—	U8 –A14
Ile	A73 , C4:G69	G34, A35, U36, A37 , A38	U12:A23, C29–G41
Met*	A73 , (G2:C71 , C3:G70) U4:A69, A5:U68	C34, A35, U36 , (C32 , U33, A37)	—
Trp	G73, A1:U72 , G2:C71, G3:C70	C34, C35, A36	—
Asp	G73, G2:C71	G34, U35, C36 , C38	G10
Lys	A73	U34, U35, U36	—
Gly	U73, G1:C72, C2:G71, G3:C70	C35, C36	—
Arg	—	C35, G36	A20
His	G–1, C73	—	—
Ser	G73, C72, G2:C71, A3:U70, R4–Y69	—	C11:G24, variable loop
Thr	G1:C72 , C2:G71	G34, G35, U36	—
Pro	A73, G72	G35, G36	G15:C48
Glu	G1:C72, U2:A71	U34 , U35, A37	U11–A24 , U13:G22:A46, Δ 47
Tyr	A73	U35	—
Cys	U73 , G2:C71, C3:G70	G34, C35, A36	G15:G48, A13:A22
Asn	G73	G34, U35, U36	—
Ala	A73, G2:C71, G3:U70 , G4:C69	—	G20
Gln	G73, U1:A72 , G2:C71, G3:C70	Y34, U35, G36 , A37, U38	G10

E. coli identity elements are as described in Giegé et al. (1998). The corresponding nucleotides present in the consensus mt tRNAs (Fig. 2) are indicated in bold and underlined. tRNAs are listed as in mt genomes, except that isoacceptors for tRNA^{Leu} and tRNA^{Ser} have been omitted.

*Nucleotides in parentheses are those additionally required in initiator tRNA^{Met}. U4:A69, A5:U68 are specific elements in elongator tRNA^{Met}.

E. coli tRNAs necessary for specific aminoacylation by the *E. coli* tyrosyl- and asparaginyl-tRNA synthetases, respectively, are present in the mitochondrial tRNAs. In most of the other cases, at least some elements are present. Within those, the discriminator base (residue 73) that is an identity element in 18 *E. coli* systems is highly conserved (and of identical nature as in *E. coli*) in 12 mt systems.

A number of other *E. coli* identity elements are found in the corresponding mitochondrial systems investigated here. Noteworthy are the anticodon triplets in tRNA^{Trp} different from *E. coli* and from mitochondria. For this system, potential identity elements within the anticodon triplet will be different in the two kingdoms. Another striking case concerns the alanine systems. A G-U base pair at nt 3–70 was shown to be the major identity element for alanylation in all known tRNA^{Ala} sequences from prokaryotes, archae, eukaryote cytoplasm, chloroplasts, and plant mitochondria (Chihade et al., 1998; Sprinzl et al., 1998). This is in combination with minor elements in the acceptor stem, namely G1–C72, G2–C71, and the discriminator base A73. Interestingly, the mt alanyl-tRNA synthetase from *Caenorhabditis elegans* has been cloned and its specificity examined (Chihade et al., 1998) and it was shown that the G3–U70 pair is retained as a positive identity signal, whereas the other elements are not conserved in nematodes. In mammalia the situation seems different,

because tRNA^{Ala} have a G3–C70 base pair, with the one exception of the platypus, *Ornithorhynchus anatinus*. One sequence (the opossum *Didelphus virginiae*) completely lacks any G-T pair in the acceptor stem. This is even more striking, because there is an average of four G-T base pairs in the secondary structure of tRNA^{Ala}, one of which is situated in the acceptor stem. Considering that other mammalian mt tRNAs (Glu, Cys, and Pro) show a similar content of G-T pairs, both in frequency and in position, the recognition by mammalian mt alanyl-tRNA synthetase seems to rely on different mechanisms, provided the G-T pair is not reestablished by an editing event. Interesting in this regard is the occurrence of T27–T43 in ~20% of the sequences. This mismatched base pair has been shown to partially restore alanyl-acceptance in an *E. coli* tRNA^{Ala} that lacks the major identity determinant G3–U70 (Hou & Schimmel, 1992).

The presence of some conserved identity elements between *E. coli* and mammalian mt tRNAs should only be considered as a starting point in the search of mitochondrial identity sets. Indeed, additional features such as other nucleotides, the different overall structural flexibility of the mt tRNAs due to the absence of D/T loop interactions, or variations in posttranscriptional modification patterns may contribute as well. Such variations may explain why *E. coli* aminoacyl-tRNA synthetases are unable to charge mt tRNAs (Kumazawa et al., 1991).

CONCLUSIONS AND PERSPECTIVES

The large number of recently published genomic sequences of mammalian mt tRNAs has enabled us to perform a vertical alignment of 679 sequences broken down into 20 families according to different amino acid specificity, and subsequently into two subfamilies in the case of tRNA^{Leu} and tRNA^{Ser} isoacceptors, according to codon recognition properties. The sequence pool, delimited on 31 fully sequenced mammalian mt genomes, was suitable for such alignment (only 3 sequences out of 682 have not been considered because of large structural deviations) and is rich in information. Not only did it define typical sequence and structural features for each of the 22 tRNA families, but it also allowed for an enlarged horizontal comparison of the different families and, as an outcome, for the delineation of a range of variability in structural features of mammalian mt tRNAs.

Characteristic patterns for each tRNA family include conservation, semiconservation, or nonconservation of nucleotides at each position within the secondary structure (except for parts of the D- and T-loops), conservation of secondary structural features and of potential tertiary interactions, distribution of noncanonical G-T pairs and mismatches, and a degree of variability in D- and T-loops sizes. As a general outcome the following conclusions can be drawn.

Secondary structures

1. Each tRNA family is made up of individual tRNAs sharing basically the same secondary structure (e.g., all tRNA^{Ser(AGY)} are missing the D-domain, all tRNA^{Ser(UCN)} are missing residue 8, all other tRNAs have a canonical global cloverleaf structure with size variations restricted to the D- and T-domains).
2. Within the same family, the size of D- and T-loops can vary to large extents (some families having rather conserved sizes, some having extremely large variations); concomitantly, the first neighboring base pair in the D- and T-stem may become disrupted and thus, the stems shortened.
3. Light tRNAs (transcribed from the DNA heavy strand) are rich in mismatches and poor in G-T pairs, whereas heavy tRNAs (transcribed from the DNA light chains) are G-T rich, and poor in mismatches.

Tertiary structures

4. Only four families, namely Leu(UUR), Leu(CUN), Glu, and Asn, have strong conservation of all potential classical tertiary interactions, including interaction between the D- and T-loops; the Ser(UCN) family does not have the core interactions but has D/T loop interactions.
5. Most families have only a few of the classical tertiary interactions conserved, excluding the D-/T-

loop interactions, all having at least a basic set of interactions between the D-stem and the variable region allowing a central structural core to form.

Primary sequences

6. The degree of conservation of primary sequences varies over a large range for each family, from very strongly conserved (Met, Leu(UUR), Leu(CUN), and Ile) to very poorly conserved (Asn).
7. Light tRNAs have D- and T-loops that are very low in G content (with the exception of the two tRNA^{Leu}); heavy tRNAs have D-loops which are very low in C residues.
8. The Val, Met, Leu(CUN), Pro, Ser, and Gln families have conserved mismatches or G-T pairs at specific sites, which might be recognition signals for molecular partners.
9. Overlap with *E. coli* aminoacylation identity elements designates the discriminator base (nt 73) and anticodon triplet as most likely identity positions.

In conclusion, with the exception of tRNA^{Ser(AGY)}, mammalian mt tRNAs do share a number of typical structural secondary and tertiary features of canonical tRNAs. tRNA^{Leu(UUR)}, tRNA^{Leu(CUN)}, and tRNA^{Gln} are very close to classical tRNAs, sharing not only typical secondary structural features, but also all the canonical tertiary interactions. tRNA^{Met} and tRNA^{Ile} are those tRNAs with the most impressive primary structure conservation. For these tRNAs, secondary and tertiary structures are also very much conserved, but incomplete with regard to classical tRNAs. Interestingly, these most conserved tRNA genes are clustered in the mt DNA.

Mitochondrial genomes are subject to an extremely high genetic drift. Indeed, the mutation rate for mt DNA is 5- to 100-fold higher than in nuclear DNA (see, e.g., Horai et al., 1992; Allen & Raven 1996; Pesole et al., 1999). Different structural, functional, and regulatory constraints apply to each individual tRNA gene, to ascertain efficient translation of the genome. Mitochondrial DNA shows two apparent evolutionary drifts, namely a tendency to economics and an enrichment of G and T nucleotides in the heavy strand, and correspondingly of C and A nucleotides in the light strand. At the level of tRNA genes, mutations at individual nucleotides will only be tolerated if they do not lead to loss of structural or functional information, or loss of interaction with partners even at the level of the gene (such as the mt transcription termination factor mTERF, which binds to the tRNA^{Leu(UUR)} gene). The results of our compilation delineating conserved and nonconserved features in each tRNA family discriminate the interplay between the two opposed evolutionary trends. The more conserved the tRNA, the more restrictions apply. D- and T-loops are the places where the evolutionary drift

is allowed to occur. The sizes of the loops varies with a tendency to small loops and the sequence is biased towards a pronounced low G content in light tRNAs and a low C content in heavy tRNAs. These parts of tRNA genes may be the sole domains within the complete mt genomes where evolutionary constraints do not apply, the rest of the genome having stronger restrictions such as codon maintenance. Point mutations within the human mt tRNA gene, correlated to various pathologies, likely affect positions of either structural or functional (or both) importance for the tRNA metabolic fate. The present mammalian tRNA structural database, to be updated at regular intervals on the web site <http://mamit-tRNA.u-strasbg.fr>, is expected to serve as an initial insight towards the understanding of the molecular mechanisms underlying the different pathologies.

MATERIALS AND METHODS

Assembly and composition of the sequence pools for vertical alignment

Thirty-one complete mammalian mt genomes have been retrieved from the European Molecular Biology Laboratory nucleotide sequence databank (Boore, 1999; Stoesser et al., 1999) and the 22 tRNA genes extracted and compiled. Table 1 gives the list of the compiled genomes. It has to be pointed out that the information contained in these genomes is somewhat biased because the sequences are not equally distributed among the different orders of mammals. Thus, the set contains 7 closely related hominoid genomes, 2 types of rhinoceros, whales, and seals, and only 1 type of the 19 other animal orders.

The sequences of the 22 different tRNA species have been aligned primarily on the basis of their anticodon triplet nucleotides, and adjusted manually for the other domains after visual inspection. This was of particular importance because the genes showed a large variability in length of their D- and T-loops. Three tRNA sequences have been removed from the compilation because of very large variations in regard to the other sequences. This was the case for the cysteine-specific tRNA from *Dasypus novemcinctus* (nine-banded armadillo), which, in contrast to the 30 other tRNA^{Cys} investigated, is missing the D-stem and loop and has instead a D-replacement loop, which is similar to mt tRNA^{Cys} from many reptiles (Macey et al., 1997). The second and third exceptions concern tRNA^{Lys} from *D. virginiae* (oppossum) and from *Macropus robustus* (wallaroo), which do not allow for a reasonable secondary tRNA-fold and might be pseudogenes (Janke et al., 1994, 1997). Thus, a total of 679 genes have been analyzed. Additional mammalian mitochondrial tRNA sequences available in the databases have not been taken into account here, to ensure an equivalent representation of tRNAs from each specificity in our analysis.

Sequence comparisons within each of the 22 tRNA families

Sequence comparison of the 31 sequences in each tRNA family (with the exceptions mentioned above, i.e., 30 in the

tRNA^{Cys} family, and 29 in the tRNA^{Lys} family) needed manual analysis so as to overcome large size variability, especially in the D- and T-regions. This variability was already observed by many authors during mt genome sequencing (see references in Boore, 1999). Once this problem was solved, simple alignments in the acceptor stems, connectors 1 (residues 8 and 9), the D-stems, connectors 2 (residue 26), the anticodon stems and loops, the variable regions, and the T-stems were possible. The CCA 3' end, typical of tRNAs but not encoded by the mt genes, was not considered here. In each of the former domains, most of which are well conserved in size over the different animal species, a statistical analysis of nucleotide composition at each position was computed, as well as a statistical analysis of the different types of interactions within the stems. Thus, for each tRNA family, a two-dimensional array (matrix) was created, consisting of the full sequences and their individual nucleotides as rows and columns, respectively. For each position, the whole set of sequences was systematically scanned and the number of occurrences of every standard nucleotide of the RNA alphabet (ACGU) stored in different variables. By means of setting these results in relation to the total number of sequences, the nucleotide composition at each position was obtained. Comparing the four nucleotide percentages according to size led to a detection of various degrees of conservation and yielded an output depicting a direct instruction for the color and character code of the graphical representation. Results are summarized in Figure 2 in the form of "typical" and "consensus" tRNAs for each amino acid specificity.

Nucleotides are considered typical when present in >50% of the sequences within a given tRNA family. Different degrees of conservation above 50% (namely $50\% < x < 90\%$, $90\% \leq x < 100\%$, and $x = 100\%$) have been calculated, and conservation within purines or pyrimidines sought. Secondary structure typical features within the stems, that is, typical Watson-Crick base pairs, G-T pairs, and mismatches, have been analyzed. As mentioned above, D- and T-loops show a large variability in size within most of the tRNA families, so that detailed information about the loops was difficult to ascertain and typical nucleotides have been found only rarely.

Limitation in gene sequence analysis

In the present work, gene sequences rather than tRNA sequences are compared. This leads to two types of limitations in the present analysis. First, editing events cannot be taken into account. Editing has been described in a number of mt tRNAs including mammalia. Indeed, a case of C-to-U editing has been described in the anticodon of the tRNA^{ASP} precursor from marsupials (Janke & Pääbo, 1993; Mörl et al., 1995; Börner et al., 1996) and can be suspected for *M. robustus* (Janke et al., 1997). Also, the discriminator base (nucleotide at position 73) of certain mt tRNAs may be changed in the processing of overlapping human mt tRNA transcripts (Reichert et al., 1998). In chicken, a G is posttranscriptionally added to the 5' end of mt tRNA^{His}, which leaves a similar possibility open for mammals. Second, epigenetic phenomena comprise all posttranscriptional base modifications, which are known to possibly affect structure (Helm et al., 1998, 1999) and function (Degoul et al., 1998) of tRNAs, in mitochondria more so than in the cytosol (Börner et al., 1997). Only a few sequences at the RNA level have been reported

so far (Roe et al., 1981; Roe, 1982; Yokogawa et al., 1991; Wakita et al., 1994; Takemoto et al., 1995; Mörl et al., 1995; Brulé et al., 1998; Helm et al., 1998), which does not yet allow for a general view of patterns of modified nucleotides in mammalian mt tRNAs.

ACKNOWLEDGMENTS

This investigation was supported by Centre National de la Recherche Scientifique (CNRS), Université Louis Pasteur (Strasbourg), Association Française contre les Myopathies (AFM) and E.C grants BIO4-CT98-0189 and QLG2-CT-1999-00660. M.H. was supported by a Marie-Curie fellowship from the European Community Training and Mobility of Researchers program, H.B. by a Poste d'accueil Institut National de la Santé et de la Recherche Médicale grant, and D.G. by Österreichische Akademie der Wissenschaften. We are grateful to Peter Schuster for support and to Caroline Paulus and Nicolas Florentz for help in calculations and drawing of Figure 2. ULP Multimedia is acknowledged for help in web page preparation.

Received May 25, 2000; returned for revision
June 26, 2000; revised manuscript received July 10, 2000

REFERENCES

- Allen JF, Raven JA. 1996. Free-radical-induced mutation vs redox regulation: Costs and benefits of genes in organelles. *J Mol Evol* 42:482–492.
- Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon JC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465.
- Boore JL. 1999. Animal mitochondrial genomes. *Nucleic Acids Res* 27:1767–1780.
- Börner GV, Mörl M, Janke A, Pääbo S. 1996. RNA editing changes the identity of a mitochondrial tRNA in marsupials. *EMBO J* 15:5949–5957.
- Börner GV, Yokobori S-i, Mörl M, Dörner M, Pääbo S. 1997. RNA editing in metazoan mitochondria: Staying fit without sex. *FEBS Lett* 409:320–324.
- Brulé H, Holmes WM, Keith G, Giegé R, Florentz C. 1998. Effect of a mutation in the anticodon of human mitochondrial tRNA^{Pro} on its post-transcriptional modification pattern. *Nucleic Acids Res* 26:537–543.
- Bullard J, Cai Y-C, Spremulli LL. 2000. Expression and characterization of the human mitochondrial leucyl-tRNA synthetase. *Biochim Biophys Acta* 1490:245–258.
- Bullard J, Cai Y-C, Demeler B, Spremulli LL. 1999. Expression and characterization of a human mitochondrial phenylalanyl-tRNA synthetase. *J Mol Biol* 288:567–577.
- Cantatore P, De Benedetto C, Gadaleta G, Gallerani R, Kroon AM, Holtrop M, Lanave C, Pepe G, Quagliariello C, Saccone C, Sbisà E. 1982. The nucleotide sequences of several tRNA genes from rat mitochondria: Common features and relatedness to homologous species. *Nucleic Acids Res* 10:3279–3289.
- Chihade JW, Hayashibara K, Shiba K, Schimmel P. 1998. Strong selective pressure to use G:U to mark an RNA acceptor stem for alanine. *Biochemistry* 37:9193–9202.
- de Bruijn MHL, Klug A. 1983. A model for the tertiary structure of mammalian mitochondrial transfer RNAs lacking the entire 'dihydrouridine' loop and stem. *EMBO J* 2:1309–1321.
- de Bruijn MHL, Schreier PH, Eperon IC, Barrell BG. 1980. A mammalian mitochondrial serine transfer RNA lacking the 'dihydrouridine' loop and stem. *Nucleic Acids Res* 8:5213–5222.
- De Giorgi C, Martiradonna A, Saccone C. 1996. Evolutionary analysis of sea urchin mitochondrial tRNAs: Folding of the molecules as suggested by the non-random occurrence of nucleotides. *J Mol Evol* 30:191–199.
- Degoul F, Brulé H, Cepanec C, Helm M, Marsac C, Leroux J-P, Giegé R, Florentz C. 1998. Isoleucylation properties of native human mitochondrial tRNA^{Ile} and tRNA^{Ile} transcripts. Implications for cardiomyopathy-related point mutations (4269, 4317) in the tRNA^{Ile} gene. *Hum Mol Gen* 7:347–354.
- Dirheimer G, Keith G, Dumas P, Westhof E. 1995. Primary, secondary and tertiary structures of tRNAs. In: Söll D, RajBhandary UL, eds., *tRNA: Structure, biosynthesis, and function*. Washington, DC: ASM Press. pp 93–126.
- Florentz C, Giegé R. 1995. tRNA-like structures in viral RNAs. In: Söll D, RajBhandary UL, eds., *tRNA: Structure, biosynthesis, and function*. Washington, DC: ASM Press. pp 141–163.
- Giegé R, Puglisi JD, Florentz C. 1993. tRNA structure and aminoacylation efficiency. *Prog Nucleic Acid Res Mol Biol* 45:129–206.
- Giegé R, Sissler M, Florentz C. 1998. Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res* 26:5017–5035.
- Goto Y-i, Nonaka I, Horai S. 1990. A mutation in the tRNA^{Leu(UUR)} gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature* 348:651–653.
- Grosjean H, Benne R. 1998. *Modification and editing of RNA*. Washington, DC: ASM Press.
- Hayashi I, Kawai G, Watanabe K. 1997a. Expression of bovine mitochondrial tRNA^{Ser(GCU)} derivatives in *Escherichia coli*. *Nucleic Acids Res* 25:3503–3507.
- Hayashi I, Kawai G, Watanabe K. 1998. Higher-order structure and thermal instability of bovine mitochondrial tRNA^{Ser(UGA)} investigated by proton NMR spectroscopy. *J Mol Biol* 284:57–69.
- Hayashi I, Yokogawa T, Kawai G, Ueda T, Nishikawa K, Watanabe K. 1997b. Assignment of imino proton signals of G-C base pairs and magnesium ion binding: An NMR study of bovine mitochondrial tRNA^{Ser(GCU)} lacking the entire D-arm. *J Biochem* 121:1115–1122.
- Helm M, Brulé H, Degoul F, Cepanec C, Leroux J-P, Giegé R, Florentz C. 1998. The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA. *Nucleic Acids Res* 26:1636–1643.
- Helm M, Giegé R, Florentz C. 1999. A Watson-Crick base-pair disrupting methyl group (m¹A9) is sufficient for cloverleaf folding of human mitochondrial tRNA^{Lys}. *Biochemistry* 38:13338–13346.
- Horai S, Satta Y, Hayasaka K, Kondo R, Inoue T, Ishida T, Hayashi S, Takahata N. 1992. Man's place in hominoidae revealed by mitochondrial DNA genealogy. *J Mol Evol* 35:32–43.
- Hou YM, Schimmel P. 1992. Functional compensation of a recognition-defective transfer RNA by a distal base pair substitution. *Biochemistry* 31:10310–10314.
- Janke A, Feldmaier-Fuchs G, Thomas WK, von Haeseler A, Pääbo S. 1994. The marsupial mitochondrial genome and the evolution of placental mammals. *Genetics* 137:243–256.
- Janke A, Pääbo S. 1993. Editing of a tRNA anticodon in marsupial mitochondria changes its codon recognition. *Nucleic Acids Res* 21:1523–1525.
- Janke A, Xu X, Arnason U. 1997. The complete mitochondrial genome of the wallaroo (*Macropus robustus*) and the phylogenetic relationship among Monotremata, Marsupialia, and Eutheria. *Proc Natl Acad Sci USA* 94:1276–1281.
- Kogelnick AM, Lott MT, Brown MD, Navathe SB, Wallace DC. 1998. MITOMAP: A human mitochondrial genome database—1998 update. *Nucleic Acids Res* 26:112–115.
- Kumazawa Y, Himeno H, Miura K-I, Watanabe K. 1991. Unilateral aminoacylation specificity between bovine mitochondria and eubacteria. *J Biochem* 109:421–427.
- Kumazawa Y, Nishida M. 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J Mol Evol* 37:380–398.
- Larsson N-G, Clayton DA. 1995. Molecular genetic aspects of human mitochondrial disorders. *Annu Rev Genetics* 29:151–178.
- Lynch M. 1996. Mutation accumulation in transfer RNAs: Molecular evidence for Muller's ratchet in mitochondrial genomes. *Mol Biol Evol* 13:209–220.
- Ma L, Spremulli LL. 1995. Cloning and sequence analysis of the human mitochondrial translation initiation factor 2 cDNA. *J Biol Chem* 270:1859–1865.
- Macey JR, Larson A, Ananjeva NB, Papenfuss TJ. 1997. Replication slippage may cause parallel evolution in the secondary structures of mitochondrial transfer RNAs. *Mol Biol Evol* 14:30–39.
- McClain WH. 1995. The tRNA identity problem: Past, present and

- future. In: Söll D, RajBhandary UL, eds., *tRNA: Structure, biosynthesis, and function*. Washington, DC: ASM Press. pp 335–347.
- Moriya J, Yokogawa T, Wakita K, Ueda T, Nishikawa K, Crain PF, Hashizume T, Pomerantz SC, MacCloskey JA, Kawai G, Hayashi N, Yokoyama S, Watanabe K. 1994. A novel modified nucleoside found at the first position of the anticodon of methionine tRNA from bovine liver mitochondria. *Biochemistry* 33:2234–2239.
- Mörl M, Dörner M, Pääbo S. 1995. C to U editing and modifications during the maturation of the mitochondrial tRNA^{Asp} in marsupials. *Nucleic Acids Res* 23:3380–3384.
- Pesole G, Gissi C, De Chirico A, Saccone C. 1999. Nucleotide substitution rate of mammalian mitochondrial genomes. *J Mol Evol* 48:427–434.
- Pütz J, Puglisi JD, Florentz C, Giegé R. 1991. Identity elements for specific aminoacylation of yeast tRNA^{Asp} by cognate aspartyl-tRNA synthetase. *Science* 252:1696–1699.
- Reichert A, Rothbauer U, Mörl M. 1998. Processing and editing of overlapping tRNAs in human mitochondria. *J Biol Chem* 273:31977–31984.
- Roe BA, Wong JF-H, Chen EY, Armstrong PA. 1981. Recombinant DNA. In: Walton AG, ed. *Proceedings of the third Cleveland symposium on macromolecules*. Amsterdam: Elsevier. pp 167–176.
- Roe BA, Wong JFH, Chen EY, Armstrong PW, Stankiewicz A, Ma D-P, McDonough J. 1982. Mammalian mitochondrial tRNAs: A modified nucleotide 3' to the anticodon may modulate their codon response. In: Slonimski P, Borst P, Attardi G, eds. *Mitochondrial genes*, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 45–49.
- Schon EA, Bonilla E, DiMauro S. 1997. Mitochondrial DNA mutations and pathogenesis. *J Bioenerg Biomembr* 29:131–149.
- Söll D, RajBhandary UL, eds. 1995. *tRNA: Structure, biosynthesis, and function*. Washington, DC: ASM Press.
- Sprinzi M, Horn C, Brown M, Ioudovitch A, Steinberg S. 1998. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* 26:148–153.
- Steinberg S, Cedergren R. 1994. Structural compensation in atypical mitochondrial tRNAs. *Nature Struct Biol* 1:507–510.
- Steinberg S, Gautheret D, Cedergren R. 1994. Fitting the structurally diverse animal mitochondrial tRNAs^{Ser} to common three-dimensional constraints. *J Mol Biol* 236:982–989.
- Steinberg S, Leclerc F, Cedergren R. 1997. Structural rules and conformational compensations in the tRNA L-form. *J Mol Biol* 266:269–282.
- Stoesser G, Tul MA, Lopez R, Sterk P. 1999. The EMBL Nucleotide Sequence Database. *Nucleic Acids Res* 27:18–24.
- Takemoto C, Koike T, Yokogawa T, Benkowski L, Spemulli LL, Ueda TA, Nishikawa K, Watanabe K. 1995. The ability of bovine mitochondrial transfer RNA^{Met} to decode AUG and AUA codons. *Biochimie* 77:104–108.
- Takeuchi N, Kawakami M, Omori A, Ueda T, Spemulli LL, Watanabe K. 1998. Mammalian mitochondrial methionyl-tRNA transformylase from bovine liver. *J Biol Chem* 273:15085–15090.
- Tiranti V, Savoia A, Forti F, D'Apolito M, Centra M, Rocchi M, Zeviani M. 1997. Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPOL) by cyberscreening of the Expressed Sequence Tags database. *Hum Mol Genet* 6:615–625.
- Ueda T, Watanabe K, Ohta T. 1983. Structural analysis of bovine mitochondrial tRNA^{Ser(AGY)}. *Nucleic Acids Res Symp Ser* 12:141–144.
- Ueda T, Yotsumoto Y, Ikeda K, Watanabe K. 1992. The T-loop region of animal mitochondrial tRNA^{Ser(AGY)} is a main recognition site for homologous seryl-tRNA synthetase. *Nucleic Acids Res* 20:2217–2222.
- Wakita K, Watanabe Y-I, Yokogawa T, Kumazawa Y, Nakamura S, Ueda T, Watanabe K, Nishikawa K. 1994. Higher-order structure of bovine mitochondrial tRNA^{Phe} lacking the “conserved” GG and TΨCG sequences as inferred by enzymatic and chemical probing. *Nucleic Acids Res* 22:347–353.
- Wallace DC. 1992. Diseases of the mitochondrial DNA. *Annu Rev Biochem* 61:1175–1212.
- Wallace DC. 1999. Mitochondrial diseases in man and mouse. *Science* 283:1482–1488.
- Watanabe K, Osawa S. 1995. tRNA sequences and variations in the genetic code. In: Söll D, RajBhandary UL, eds., *tRNA: Structure, biosynthesis, and function*. Washington, DC: ASM Press. pp 225–250.
- Watanabe Y-i, Kawai G, Yokogawa T, Hayashi N, Kumazawa Y, Ueda T, Nishikawa K, Hirao I, Miura K, Watanabe K. 1994a. Higher-order structure of bovine mitochondrial tRNA^{Ser(UGA)}: Chemical modification and computer modeling. *Nucleic Acids Res* 22:5378–5384.
- Watanabe Y-i, Tsurui H, Ueda T, Furushima R, Takamiya S, Kita K, Nishikawa K, Watanabe K. 1994b. Primary and higher order structures of nematode (*Ascaris suum*) mitochondrial tRNAs lacking either the T or D stem. *J Biol Chem* 269:22902–22906.
- Westhof E, Dumas P, Moras D. 1985. Crystallographic refinement of yeast aspartic acid transfer RNA. *J Mol Biol* 184:119–145.
- Wolstenholme DR, Macfarlane JL, Okimoto R, Clary DO, Wahleithner JA. 1987. Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. *Proc Natl Acad Sci USA* 84:1324–1328.
- Wolstenholme DR, Okimoto R, Mcfarlane JL. 1994. Nucleotide correlations that suggest tertiary interactions in the TΨ-replacement loop-containing mitochondrial tRNAs of the nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Nucleic Acids Res* 22:4300–4306.
- Yokogawa T, Watanabe Y, Kumazawa Y, Ueda T, Hirao I, Miura K, Watanabe K. 1991. A novel cloverleaf structure found in mammalian mitochondrial tRNA^{Ser(UCN)}. *Nucleic Acids Res* 19:6101–6105.
- Zhang Y, Spemulli LL. 1998. Roles of residues in mammalian mitochondrial elongation factor Ts in the interaction with mitochondrial and bacterial elongation factor Tu. *J Biol Chem* 273:28142–28148.